

1 CCGCAACCCCC GACGGCGCCC CAAACGCTGT TGCGCCGCGC GCCCCGCCA
 51 GCCCCGGCCTC GCGCTGGTCC CGGTCTCGCC CCGCAGCCCT CGATCTCCG
 101 TGACTTCCTC GGCGAGGCCG CCTGCGCCCTC TGGGACCATG TTGCGCTGGC
 151 TGCAGGGACTT CGCGCTGCC ACCGCGGCCT GCCAGGACGC GGAGCAGCCG
 201 ACGCGCTACG AGACCCTCTT CCAGGCACGT GACCGCAATG GGGACGGAGT
 251 GGTGGACATC GGGCAGCTGC AGGAGGGCT CAGGAACCTG GGCATCCCTC
 301 TGGGCCAGGA CGCCGAGGAG AAAATTTTA CTACTGGAGA TGTCAACAAA
 351 GATGGGAAGC TGGAGTTGA AGAATTATG AAGTACCTTA AAGACCATGA
 401 GAAGAAAATG AAATTGGCAT TTAAGAGTTT AGACAAAAT AATGATGGAA
 451 AAATTGAGGC TTCAGAAATT GTCCAGTCTC TCCAGACACT GGGTCTGACT
 501 ATTTCTGAAC AACAAAGCAGA GTTGATTCTT CAAAGCATTG ATGTTGATGG
 551 GACAATGACA GTGGACTGGA ATGAATGGAG AGACTACTTC TTATTTAAC
 601 CTGTTACAGA CATTGAGGAA ATTATCCGT TCTGAAACAA TTCTACAGGA
 651 ATTGACATAG GGGATAGCTT AACTATTCCA GATGAATTCA CGGAAGACGA
 701 AAAAAAATCC GGACAATGGT GGAGGGCAGCT TTTGGCAGGA GGCATTGCTG
 751 GTGCTGTCTC TCGAACAAAGC ACTGCCCTT TGGACCGTCT GAAAATCATG
 801 ATGCAGGGTTC ACGGTTCAAA ATCAGACAAA ATGAACATAT TTGGTGGCTT
 851 TCGACAGATG GTAAAAGAAG GAGGTATCCG CTCGTTTGG AGGGGAAATG
 901 GTACAAACGT CATCAAATTG GCTCCTGAGA CAGCTGTTAA ATTCTGGCA
 951 TATGAACAGT ACAAGAAGTT ACTTACTGAA GAAGGACAAA AAATAGGAAC
 1001 ATTTGAGAGA TTTATTCTG GTTCCATGGC TGGAGCAACT GCACAGACTT
 1051 TTATATATCC AATGGAGGTT ATGAAAACCA GGCTGGCTGT AGGCAAAACT
 1101 GGGCAGTACT CTGGAATATA TGATTGTCG AAGAAGATT TGAAACATGA
 1151 AGGCTTGGGA GCTTTTACA AAGGCTATGT TCCCAATTAA TTAGGTATCA
 1201 TACCTTATGC AGGCATAGAT CTTGCTGTG ATGAGCTCTT GAAGTCCTAT
 1251 TGGCTGGATA ATTTGCAAA AGATTCTGTA AACCCCTGGAG TCATGGTGT
 1301 GCTGGGATGC GGTGCCTTAT CCAGCACCTG TGGTCAGCTG GCCAGCTACC
 1351 CATTGGCTTT GGTGAGAACT CGCATGCAGG CTCAGCCAT GTTAGAAGGT
 1401 TCCCCACAGC TGAATATGGT TGGCCTCTT CGACGAATTAA TTTCCAAAGA
 1451 AGGAATACCA GGACTTTACA GAGGCATCAC CCCAAACTTC ATGAAGGTGC
 1501 TCCCTGCTGT AGGCATCACT TATGTTTTT ATGAAAATAT GAAGCAAAC
 1551 TTAGGAGTAA CCCAGAAATG ATGTTGCATT TTTTGCTTTA GCCTGATAAT
 1601 TGAAACTTTC AACAAATCTCT GGAGTGACTT TTTCTCTCG AATTGAAACA
 1651 AGTCTATGGC AAAAGAAGCT GCATTTTTT CACAAAAGGG AAGACGGTAA
 1701 CAATGGTCAC TTCAAACTTT TGGGCTAAAT TATATGTACA CAGAAATGTT
 1751 CAAAATCATA GTTTTAATGT GTTTTGAAAA GGCCACACAA TTATACTTTA
 1801 TCTTTCTTA ATAATCCTGC AAATCTCTGC CCTGAATCCG AAATCTGAAA
 1851 ATGTAATGGC TTGAACAAAA TTTGTTTTGT GTGTTAGAGT TATAAATCAT
 1901 TAATCTTAT TTGGGTGGT TTACGTTAT GCCAGTTCT TTATATTTA
 1951 ATTTCTTGT TTATATATTG TGAATGTCTT TATAGATTTC TTTAAATTTC
 2001 CTTATAGAAC CATTAAATAGA AAATCATTAC ATTTAAAATA TACCTTACAG
 2051 CAAAAGCATC CAAATAAGTA TAGGGTTTAT GTCCTTATTT TTCTTCAGC
 2101 TGAATACGAA TGAACACAGT GGTGGAATTG CTGAAGGGAA GTGATGAAAT
 2151 TATATTTATT TCAGTGGGCA CTTTTCCATT TTACCACTGT ACCATTATT
 2201 GGTTCTGGA GTTATACACT AATTTTCAGT ATATTACTGT TAAATTACCA
 2251 ACACAAGGC AATTATTTGA AAGATTCCGT TTATCCTGCC ATTGCTTGT
 2301 AAAGCAGCAG GAAACGAAAT TTTTGACTT GTATCAGCTT CTGCAGAGCA
 2351 TCTTTGTTTT CCTTTGTCCT TTGTTCTTA CCTTTGAAT CAGATCCGT
 2401 TTTAGTCAGG AAGACTTCTT GGGACCATTC TTAGTAACCT GAAATTCTT
 2451 TTTTAATTGC ATGAAGTGGA TTGATCATGA GCAAGTGTG GGCTTTATT
 2501 CTCCCTCACT GGTGAATATC CTTTGAACCTT GCTGTTGCA ATATGGCAG
 2551 CCACAAAGGG GGAGAGATGC CTATTAATC GGCGGGGTGT ATGACTTCTG
 2601 AAAACATTGG ATACCCTATT TTGAAAAGGG AAAGGCCAA TTTGGGGAAA
 2651 CATATACCAA TGCATGATTG CTG

FEATURES:

5'UTR: 1-137
 Start Codon: 138
 Stop Codon: 1569
 3'UTR: 1572

HOMOLOGOUS PROTEINS:

Top BLAST Hits:

	Score	E
CRA 335001098641184 /altid=gi 11360341 /def=pir T50686 peroxis...	927	0.0
CRA 11000479457833 /altid=gi 6841066 /def=gb AAF28888.1 AF12330...	834	0.0
CRA 18000005183605 /altid=gi 7504235 /def=pir T22688 hypotheti...	432	e-120
CRA 1000682325160 /altid=gi 7499323 /def=pir T21074 hypothetic...	377	e-103
CRA 89000000196990 /altid=gi 7294582 /def=gb AAF49922.1 (AE003...	348	9e-95
CRA 150000075553401 /altid=gi 9758252 /def=dbj BAB08751.1 (AB0...	339	5e-92
CRA 335001098657884 /altid=gi 11358611 /def=pir T49871 peroxis...	330	2e-89
CRA 163000046661776 /altid=gi 10176874 /def=dbj BAB10081.1 (AB...	326	4e-88
CRA 105000014652720 /altid=gi 10798831 /def=dbj BAB16462.1 (AP...	200	3e-50
CRA 335001098655048 /altid=gi 11277065 /def=pir T47703 Ca-depe...	199	6e-50

BLAST dbEST hits:

gi 10145202 /dataset=dbest /taxon=96...	1108	0.0
gi 1437155 /dataset=dbest /taxon=9606 ...	801	0.0
gi 10333851 /dataset=dbest /taxon=96...	745	0.0
gi 8469752 /dataset=dbest /taxon=960...	363	8e-98
gi 11684041 /dataset=dbest /taxon=96...	307	4e-81

EXPRESSION INFORMATION FOR MODULATORY USE:

library source:

Expression information from BLAST dbEST hits:

gi|10145202 Placenta Choriocarcinoma
gi|1437155 Retina
gi|10333851 Uterus leiomyosarcoma
gi|8469752 Breast
gi|11684041 Ovary fibrotheoma

Expression information from PCR-based tissue screening panels:

Leukocyte

1 MLRWLRDFAL PTAACQDAEQ PTRYETLFQA LDRNGDGVV D IGELOEGLRN
51 LGIPLGQDAE EKJFTTGDVN KDGKLDFFEF MKYLDHEKK MKLAFKSLDK
101 NNDGKIEASE IVQSLQTLGL TISEQQAEI LQSIDVDGTM TVDWNEWRDY
151 FLFNPVTDIE EIIRFWKHST GIDIGDSLTI PDEFTEDEKK SGQWWRQLLA
201 GGIAGAVSRT STAPLDRLKI MMQVHGSKSD KMNIFGGFRQ MVKEGGIRSL
251 WRGNNTNVIK IAPETAVKFW AYEQYKKLLT EEGQKIGTFE RFISGSMAGA
301 TAQTFIYPME VMKTRLAVGK TGQYSGIYDC AKKILKHEGL GAFYKGYVPN
351 LLGIIPYAGI DLAVYELLKS YWLNDNAFKDS VNPGVMVLLG CGALSSTCGQ
401 LASYPLALVR TRMQAQAMLE GSPQLNMVGL FRRISKEGI PGLYRGITPN
451 FMKVLPNAVGI SYVYYENMKQ TLGVTQK

FEATURES:

Functional domains and key regions:

[1] PDO000001 PS000001 ASN_GLYCOSYLATION
N-glycosylation site

254-257 NGTN

[2] PDO000005 PS000005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site

Number of matches: 2

1	229-231	SDK
2	475-477	TQK

[3] PDO000006 PS000006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site

Number of matches: 8

1	22-25	TRYE
2	65-68	TTGD
3	121-124	TISE
4	157-160	TDIE
5	170-173	TGID
6	179-182	TIPD
7	185-188	TEDE
8	227-230	SKSD

[4] PDOC00008 PS00008 MYRISTYL
N-myristylation site

Number of matches: 16

1	52-57	GIPLGQ
2	119-124	GLTISE
3	171-176	GIDIGD
4	201-206	GGIAGA
5	202-207	GIAGAV
6	245-250	GGIRSL
7	253-258	GNGTNV
8	283-288	GQKIGT
9	295-300	GSMAGA
10	322-327	GQYSQI
11	326-331	GIYDCA
12	359-364	GIDLAV
13	392-397	GALSST
14	399-404	GQLASY
15	442-447	GLYRGI
16	446-451	GITPNF

[5] PDOC00018 PS00018 EF HAND
EF-hand calcium-binding domain

Number of matches: 3

1	32-44	DRNGDGVVDIGEL
2	68-80	DVNKGKLDFEFF
3	99-111	DKNNNDGKIEASEI

Membrane spanning structure and domains:

Helix	Begin	End	Score	Certainty
1	292	312	1.053	Certain
2	345	365	0.613	Putative
3	381	401	1.544	Certain
4	446	466	0.733	Putative

BLAST Alignment to Top Hit:

>CRA|335001098641184 /altid=gi|11360341 /def=pir|T50686 peroxisomal
Ca-dependent solute carrier [imported] - rabbit
/org=rabbit /taxon=9986 /dataset=nraa /length=475
Length = 475

Score = 927 bits (2371), Expect = 0.0
Identities = 454/477 (95%), Positives = 466/477 (97%), Gaps = 2/477 (0%)

Query: 1 MLRWLRDFALPTAACQDAEQPTRYETLFQALDRNGDGVVDIGELQEGLRNLGIPLGQDAE 60
MLRWLR F LPTAACQ AE PTRYETLFQALDRNGDGVVDI ELQEGL++LGIPLGQDAE
Sbjct: 1 MLRWLRGFVLPPTAACQGAEPPTTRYETLFQALDRNGDGVVDIRELQEGLKSLGIPLGQDAE 60

Query: 61 EKIFTTGDVNKGKDLDFFFMKYLKDHEKKMKLAFKSLDKNNNDGKIEASEIVQLTLGL 120
EKIFTTGDVNKGKDLDFFFMKYLKDHEKKMKLAFKSLDKNNNDGKIEASEIVQLTLGL
Sbjct: 61 EKIFTTGDVNKGKDLDFFFMKYLKDHEKKMKLAFKSLDKNNNDGKIEASEIVQLTLGL 120

Query: 121 TISEQQAEILILQSIDVDGTMVDWNEWRDYFLFNPVTDIEEIIRFWKHSTGIDIGDSLTI 180
TISEQQAEILILQSID DGTMTVDWNEWRDYFLFNPV DIEEIIRFWKHSTGIDIGDSLTI
Sbjct: 121 TISEQQAEILILQSIDADGTMVDWNEWRDYFLFNPVADIEEIIRFWKHSTGIDIGDSLTI 180

Query: 181 PDEFTEDEKKSGQWWRQLLAGGIAGAVSRTSTAPLDRLKIMMQVHGSKSDKMNIFFGGFRQ 240
PDEFTE+E+KSGQWWRQLLAGGIAGAVSRTSTAPLDRLK+MMQVHGSKS MNIFGGFRQ
Sbjct: 181 PDEFTEERKSGQWWRQLLAGGIAGAVSRTSTAPLDRLKVMMQVHGSKS--MNIFGGFRQ 238

Query: 241 MVKEGGIRSLWRGNNTNVVIKIAPETAVKFWAYEQYKKLLTEEGQKIGTFERFISGSMAGA 300
M+KEGG+RSLWRGNNTNVVIKIAPETAVKFW YEQYKKLLTEEGQKIGTFERFISGSMAGA
Sbjct: 239 MIKEGGVRSLSWRGNNTNVVIKIAPETAVKFVWYEQYKKLLTEEGQKIGTFERFISGSMAGA 298

Query: 301 TAQTFIYPMEVMKTRLAVGKTGQYSGIYDCAKKILKHEGLGAFYKGYVPNLLGIIPYAGI 360
TAQTFIYPMEVMKTRLAVGKTGQYSGIYDCAKKILK+EG GAFYKGYVPNLLGIIPYAGI
Sbjct: 299 TAQTFIYPMEVMKTRLAVGKTGQYSGIYDCAKKILKYEFGFAFYKGYVPNLLGIIPYAGI 358

Query: 361 DLAVYELLKSYWLDNFAKDSVNPVMVLLGCGALSSTCGQLASYPLALVRTRMQAQAMLE 420
DLAVYELLKS+WLDNFAKDSVNPVG+VLLGCGALSSTCGQLASYPLALVRTRMQAQAMLE
Sbjct: 359 DLAVYELLKSHWLDNFAKDSVNPVGVLVLLGCGALSSTCGQLASYPLALVRTRMQAQAMLE 418

Query: 421 GSPQLNMVGLFRRRIISKEGIPGLYRGITPNFMKVLPAVGISYVYYENMKQTLGVTQK 477
G+PQLNMVGLFRRRIISKEG+PGLYRGITPNFMKVLPAVGISYVYYENMKQTLGVTQK
Sbjct: 419 GAPQLNMVGLFRRRIISKEGLPGLYRGITPNFMKVLPAVGISYVYYENMKQTLGVTQK 475

>CRA|11000479457833 /altid=gi|6841066 /def=gb|AAF28888.1|AF123303_1
(AF123303) calcium-binding transporter [Homo sapiens]
/org=Homo sapiens /taxon=9606 /dataset=nraa /length=411
Length = 411

Score = 834 bits (2132), Expect = 0.0
Identities = 409/410 (99%), Positives = 409/410 (99%)

Query: 8 FALPTAACQDAEQPTRYETLFQALDRNGDGVVDIGELQEGLRNLGIPLGQDAEEKIFTTG 67
F LPTAACQDAEQPTRYETLFQALDRNGDGVVDIGELQEGLRNLGIPLGQDAEEKIFTTG
Sbjct: 1 FVLPTAACQDAEQPTRYETLFQALDRNGDGVVDIGELQEGLRNLGIPLGQDAEEKIFTTG 60

Query: 68 DVNKDGKLDFFFMKYLKDHEKKMKLAFKSLDKNNNDGKIEASEIVQLTLGLTISEQQA 127
DVNKDGKLDFFFMKYLKDHEKKMKLAFKSLDKNNNDGKIEASEIVQLTLGLTISEQQA
Sbjct: 61 DVNKDGKLDFFFMKYLKDHEKKMKLAFKSLDKNNNDGKIEASEIVQLTLGLTISEQQA 120

Query: 128 ELILQSIDVDGTMVDWNEWRDYFLFNPVTDIEEIIRFWKHSTGIDIGDSLTI PDEFTE 187
ELILQSIDVDGTMVDWNEWRDYFLFNPVTDIEEIIRFWKHSTGIDIGDSLTI PDEFTE
Sbjct: 121 ELILQSIDVDGTMVDWNEWRDYFLFNPVTDIEEIIRFWKHSTGIDIGDSLTI PDEFTE 180

Query: 188 EKKSGQWWRQLLAGGIAGAVSRTSTAPLDRLKIMMQVHGSKSDKMNIFFGGFRQMVKEGGI 247

EKKSGQWWRQLLAGGIAGAVSRTSTAPLDRLKIMMQVHGSKSDKMNI FGGFRQMVKEGGI
 Sbjct: 181 EKKSGQWWRQLLAGGIAGAVSRTSTAPLDRLKIMMQVHGSKSDKMNI FGGFRQMVKEGGI 240

 Query: 248 RSLWRGNGTNVIKIAPETAVKFWAYEQQYKKLLTEEGQKIGTFERFISGSMAGATAQTFIY 307
 RSLWRGNGTNVIKIAPETAVKFWAYEQQYKKLLTEEGQKIGTFERFISGSMAGATAQTFIY
 Sbjct: 241 RSLWRGNGTNVIKIAPETAVKFWAYEQQYKKLLTEEGQKIGTFERFISGSMAGATAQTFIY 300

 Query: 308 PMEVMKTRLAvgKTGQYSIYDCAKKILKHEGLGAFYKGYVPNLLGIIPYAGIDLAVYEL 367
 PMEVMKTRLAvgKTGQYSIYDCAKKILKHEGLGAFYKGYVPNLLGIIPYAGIDLAVYEL
 Sbjct: 301 PMEVMKTRLAvgKTGQYSIYDCAKKILKHEGLGAFYKGYVPNLLGIIPYAGIDLAVYEL 360

 Query: 368 LKSYWLDNFAKDSVNPGMVLLCGALSSTCGQLASYPLALVRTRMQAQA 417
 LKSYWLDNFAKDSVNPGMVLLCGALSSTCGQLASYPLALVRTRMQAQA
 Sbjct: 361 LKSYWLDNFAKDSVNPGMVLLCGALSSTCGQLASYPLALVRTRMQAQA 410

Score = 80.0 bits (194), Expect = 6e-14
 Identities = 80/388 (20%), Positives = 156/388 (39%), Gaps = 59/388 (15%)

Query: 95 FKSLDKNNNDGKIEASEIIVQLQTLGLTISEQQAELILQSIDV--DGTMTVDWNEWRDYFL 152
 F++LD+N DG ++ E+ + L+ LG+ + + E I + DV DG +
 Sbjct: 21 FQALDRNGDVVDIGELQEGLRNLGIPLGQDAEKKIFTTGDVNKDGLK----- 68

 Query: 153 FNPVTDIEIIIRFWKHSTGIDIGDSLTIPEFTEDEKSGQWWRQLLAGGIAGAVSRTST 212
 D EE +++ K + EKK ++ L +
 Sbjct: 69 -----DFEEFMKYLK-----DHEKKMKLAFKSLDKNNNDGKIEASEIV 105

 Query: 213 APLDRLKIMMQVHGSKSDKMNI FGGFRQMVKEGGIRSLWRGNGTNVIKIAPETAVKFWAY 272
 L L + + ++ +I V R + N I E ++FW +
 Sbjct: 106 QSLQTLGLTISEQQAELILQSIDVDGTMVDWNEWRDYFLFPVTDI---EEIIRFWKH 161

 Query: 273 EQYKKL-----LTEEGQKIGTFER-FISGSMAGATAQTFIYPMEVMKTRLAV-GKT 321
 + TE+ +K G + R ++G +AGA ++T P++ +K + V G
 Sbjct: 162 STGIDIGDSLTIPEFTEDEKSGQWWRQLLAGGIAGAVSRTSTAPLDRLKIMMQVHGSK 221

 Query: 322 GQYSIYDCAKKILKHEGLGAFYKGYVPNLLGIIPYAGIDLAVYELLKSYWLDNFAKDSV 381
 I+ +++K G+ + ++G N++ I P + YE K ++
 Sbjct: 222 SDKMNI FGGFRQMVKEGGIRSLWRGNGTNVIKIAPETAVKFWAYEQYKKL---LTEEGQ 277

 Query: 382 NPGVMVLLCGALSSTCGQLASYPLALVRTRMQAQMLEGSPQLNMVGLFRRIISKEGIP 441
 G G+++ Q YP+ ++TR+ A+ + + ++I+ EG+
 Sbjct: 278 KIGTFERFISGSMAGATAQTFIYPMEVMKTRL---AVGKTGQYSIYDCAKKILKHEGLG 334

 Query: 442 GLYRGITPNFMKVLPAVGISYVVYENMK 469
 Y+G PN + ++P GI VYE +K
 Sbjct: 335 AFYKGYVPNLLGIIPYAGIDLAVYELLK 362

Hmmmer search results (Pfam):

Model	Description	Score	E-value	N
PF00153	Mitochondrial carrier proteins	305.4	3e-88	1
PF00036	EF hand	50.7	1.7e-12	3
PF00404	Dockerin domain type I	9.7	0.26	1
PF01978	Protein of unknown function	2.7	9.5	1

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF00036	1/3	27	51 ..	5	29 .]	18.7	0.002
PF00404	1/1	67	85 ..	1	22 []	9.7	0.26
PF00036	2/3	61	87 ..	3	29 .]	19.7	0.001
PF00036	3/3	90	118 ..	1	29 []	17.2	0.0051
PF01978	1/1	110	121 ..	1	13 [.	2.7	9.5
PF00153	1/1	193	472 ..	1	313 []	305.4	3e-88

1 AACCCATGTT AGTGTGCAGT TCTGCTGGCA CACACATGCA GTTGTGTAAC
 51 CACTACCACC AAAAGCAAGA TGAAAATAG CTCCATCACC CCCACAAGCC
 101 TTCTGATGCT CTTTGTCAT CAATTCCCTT CCCGCTAGTC ACAACTGGTA
 151 ACTACTGATT TGTTTCTGT CCCTATAAGTT TTGCCTTTTC CAGAATGTCA
 201 TTGTTGACAG GTATCAGTAA TTCATTCTT TTTATTGCTA ATTACTATCT
 251 CACTGTATGA ATGCAACACA GGTTGTTAAC CAGTCACCC GTTAAAGAAC
 301 ATTTTGTTC TGCGCTTGAC AGTTATGAAT AGAACTGCTA TAAACCTCA
 351 AGTAAAAGTT TTGGTGTGAA GATAATTTC TCAGCAAAAA CGCTGACAGG
 401 TAATTTTCT AAGTATTACT TTTTTAAAAA AGTAAAATAG CCTGTAGCCC
 451 CAGCTACTCA GGAGGCTGAG GCAGGAGAAT AGCTGAACC CAGGAGGCAG
 501 AGGTTGCGAGT GAGTTGAGAT TGTGCCACTG CATTCCAGCC TGGGCGACAG
 551 AGCTAGACTG TCTCAAAGAA AAAAAAAA AATAACAAAT AAATAAAAAG
 601 TAAATGAAA GCATGTAAGT GTAAGATGAC TAGTCAAGC AACCTCTCTT
 651 CAACTACAGA GTATTCAAGAG TAGAGATTAA AAGAGGTTTT CAAGGACAGA
 701 GAAAATTGAG AGTTGAAGG CAGTTCCAAA GGAAGGCAAT GATTCTTAAT
 751 AAGACTGGAA GTTGGAAAGTA ATATAAAAAG ATAATCAGT TTCAAGATGA
 801 TTTTACTAAG CAGGCAGGCC TTAATTACAA AATTCTAGAT TCATACATAT
 851 CTTAACACATA CAAAATGATA TGAGGAGGAGG TAAGTTCAAGG GTCTGAGTTC
 901 CTGGCTGTT TGGAACTGAA TTTCTGTTGTA GTGATTCAAGA AGATGTGAGA
 951 CACCCAATT TACAAGTACA GAGGTATCTT CTTTCTGCA AACAGCAGTA
 1001 CAACAATAGT TCCTCTTACG CAGCTGTGAA TGAACAGGAT TATTACAATT
 1051 AATGATATCT CATTGATTG GCGCCTTAGA GAATTAAGAC CTTTCACACC
 1101 TAATATACAA CTTTGTGTTG AAGGCAGATA TTTATATTCT CATTTACTG
 1151 ATGAGAGACT ACCCGGAGAC GCTATGTCAC ACCTGAAGGA TTAGGTACTT
 1201 TCTCTGTTAA GTCCAATGTT CTTCCGTTA TTCCATGCTA GGCAGTAATA
 1251 AGTTCTGTC TGCCTGAGTA ATAAGCTCCA AACCTCGGAA CTGCACCCAT
 1301 CTTGAGAAGG AGGAGGGCGC TGTGGTTTT TCTGATAAGT GCAGCTGGCA
 1351 GACACTCTAT ACCGCTTAATC ACGGGCAAT CCTACCTAAG CTGCCTACCA
 1401 AACTAGTCCT TCTTTTCCCC GTTGCCACG CAGATGGCTG TTGATCTTT
 1451 CTGCAACAAA TCCAGGAGTT TCTCCTTTTT GTTTATAAT TGCTCCAATA
 1501 GATGCTTTAG GATTTAACTC TCTGCTTTTT AAAGCAGAAT CGCCATCCCA
 1551 GGTGTGCAAC CACGAAAAAA TTAGACATCC GTGAGAGACA ATGCCCTCCA
 1601 TGGCCAGTT TCCAGGCAGA GAGAAGCAGC TCTGGGCTGA CGGCCAAGGC
 1651 TCCGGCCCGA GAGGGTCTTT AAGTGGAGTA ACCAGTCTTC AAGACCCCGC
 1701 TCCCAAGCCA CCGACGCGCT GACGCTGCAG CCCTGGACCT GCTGGGGGCC
 1751 TCTTCCTCGG ACCCGCATGC TGACAGCGGG ACTGGCAACT GGGCAGAGGT
 1801 CGACCCCGGG TCCGCACAGC ACCTCCCCAG ACCCAGCTCC CAGCTCCCTC
 1851 ACTTCCGGCT CTCTGGAGGC GGGCCCGGCC AGTGGCCCG AGGCCAGCGC
 1901 GGGAGCTCC TCCCCAGCAG CGGGGGAGC GCCCACCCCT GCGCGCCCGC
 1951 CGGGCTCGGG TGGGGCTCC GCTCCTGCAGC CCTGCGCGCC GCAGCCGCAC
 2001 CCCCCGACGGC GCCCCAAACG CTGTTGCAGC CGGCCGCCCCG CCCAGCCCCG
 2051 CCTCGCGCTG GTCCCCGGTCT CGCCCCCGCAG CCCTCGATCT CCCGTGACTT
 2101 CCTCGGCCAG GCCGCCTGCG CCTCTGGAC CATGTTGCAGC TGGCTGCGGG
 2151 ACTTCGTGCT GCCCACCGCG GCCTGCCAGG ACGGGAGCA GCGCACGCC
 2201 TACGAGACCC TCTTCCAGGC ACTGGACCGC AATGGGGACG GAGTGGTGGA
 2251 CATCGCGAG CTGCAGGAGG GGCTCAGGAA CCTGGGCATC CCTCTGGGCC
 2301 AGGACGCCGA GGAGGTGGGT CGCCGCCGGG GCGCCGCGCTG AGCGTAGGG
 2351 GGGCTCGGG CGCTGGGAGC ACTGCGAGGA CCGAGGAGGG CGGCGGCTTG
 2401 AGGCAGTTGCC AGGAGAGGAA GGAGGAACGT TGGCGCCAG CGCTCCGGTG
 2451 GCTTCAGAAA CTCGGCGTG GGGCCGCGAC CGGCAGACCC GGTAAACAGAA
 2501 GTGGGTCTATA ATACGAAAGT CTACTGGTAT TTGTCAGAT AAAATGAGTG
 2551 TTGTGGACAC TCTGGCCAC GGGCACTGTT AAATTTTAA GACACTTTG
 2601 TCCTGAATCC ATCCCAGGTT CTTTGTTC TGTGTTAATA CCTTGAGAC
 2651 ATGTAATCCG TTTTAGCTGT CAGACTTCAG TGGGTCCCAA GTTTTGATA
 2701 AAGGCGCACA CATTCGATCT CTTTCAAGC TGCTTGTAA CAGCAGCTAT
 2751 GTGTATTGTC TACTGTTGA AAACGTGTT AAAACCAATC GCGTGTTTCC
 2801 CCCACTTCCT GTGAGAAGG AATGGCGGC TTCCATTGTT TAAGACATTG
 2851 CTAGGTTAAT GCCCTAGGTA CATAAAATGAA TCTGAAGGGT TGACTTGACC
 2901 TGGCAGCTGAG CAAATTCTATT TTCTCTGAGT CATCTTAACG GTGCCCTGA
 2951 ACTTCGCCC CTTTAGTAGG GTGGAGATAT GTGGAACCTTC TCCAACCCCTG
 3001 TTGAAGCGTT CCCTGACACT GGCATTCTCT TATCCAAAGA GGGAAAGTGA
 3051 TTAGGTTACT ATGAGGGCCA ACAACTGTTA TATAGTTATA TTTCACCTCT
 3101 CTTTAAATGTT CTTGGTAGT TATAGGCCTC TTCAGTTAC TGTTTCTTCT

3151 AGAGTCAGAT TTAGTAAGTT ACAATTCTT TTGAAACTGC CTGTTCTGTC
3201 CAAGGTTCAT AATACTCACC GATGATTTA TAACACTTCT GACTGAATCT
3251 GTAGGTAGGT TCTCTATTTC ATTCCCTCATA TCTATCCCTT TCTCCCCCTTC
3301 AATCTTGCCA AAGTTTGTTG TATTTTATTC ATACTTGAA GGAACCAACT
3351 TTTGGTACTT TGTGCTGATT GTCCCCAGAAA TGCCCCAGTT GGAGTTCCCC
3401 ACCATGTCCA ATCATTGGCT GGAAGCAGCC CAGGAAAGGG ACGACCTTGC
3451 TGCAGTGCAT CAGCAGATGC CAGGGTTAGA GGCTAGAGAG TGGAAGTCAA
3501 CTGTGTCCCT CACAGTAGGT GCCTTGAAG GGAGATCTCA GTGGTACAAC
3551 TCCATGGTCC CTACAAAT CAAAAGCTCT TTGGAGTGCT CAATGATTTT
3601 TAAGATGTA AAGGGATCCT GAGATCAAAA AGCTTGAGAA TTGCTGCTGT
3651 ATCACCACTTT TTACGTAATT GCATCATATT CTGTTATATG TTGTTGTCAT
3701 AGTATATGTT ACCAATTCTT TTTAAATCAC CTTTTACTTT ATTGATAGTT
3751 TAAAAACGAT TGTAAGTCAA ATTGCAATGG ATGTCCTTG TATTCAATT
3801 CTCATTCTGG TCCAGTTACT TTCTGTAGGAT AAATTTGAG GAGTGGACAT
3851 TGCTGAGTCT GAAGGTAACA CACATTTAA ACTGGGATAC GTATTGCCCT
3901 TCGGAAACCT TAGACCCATT TTCACCTTT TGACTGACAG TGCTTGCTTC
3951 TCCACATCCT CGCTCATTCA GGGTATCAGT CTTTGTAAAG TCTCCTATT
4001 TGCAGGTGAA ATTCCCTTTT ATTTCCTGTC TTAGTCCATT TAGTGGTGC
4051 ATAGTGAAT ATCTGAGACA GGGTAATTAA TAAAGAAAAG ACATTTATT
4101 AGCTCACAGT TCCGCAGGCT GGGAAAGTTA AGAACGCGTGG TGCTGGCATIC
4151 TGCTGGACTC CTGGGGAGGG CTTTCCTGCT GTGTACAAC ATGGTGGAAA
4201 GTCAAAGTGG AAGTGGACAT GTGTGAAGAA GCAAAATCCG AGGGGTGTCC
4251 TGGCTTTATA GCAACCCAGC CTCGAGGGAA CTGATCCATT ACTGAGGGAA
4301 CTAATTCACT CTCATGAGAG AGAGAACTCA CTCACTACTG CAAGAATGAC
4351 ACCAAGCCAT TCATGAGGGT TCTGCCTCCG TAACCCCTGAC ACCTCCTGCT
4401 AGGTCCCTCC TCCCAACACG GCCACATCAG GGATCAGACT TCAACATGAG
4451 TTTTGTGGG GACAAACAAA ACGTAGCACT TGCTTGCCT TTGGTTCTA
4501 TTCACATCCT CCACAGGATT GCATTATGCC TACCCATTG GTGAGGGCAG
4551 TCTTCTTTAA TTGGTTACT GATTCAAATG CTACCCCTCCT CCAGAGACAT
4601 CCTCACAGAC ACACCCAGAA ATCATGTTT ACCAGTTATC TGGGCATCCC
4651 TTAGTCCAGA CGAGTTGATA CATAAAATTA ACCATCACAC ATGGGATAGA
4701 ATTAGGATTA CACAGTCAAC CTTTATGGG GAAAATTTCA GAGGCATGTC
4751 AGGGTTTAT GTAATGTCAA GGAGTGAGGA CATTGGCTAC TTGAGCATAG
4801 AAATGAGAAC TGTGGGTGA CTCTCGGTG GAAAGTTCA AGGTAGTAGT
4851 TTGTATCTAA GCCAATACT CAGCTTGAAAG CAAAATCTCT ATAAATTTC
4901 ATCTGATTTG ATCTCATCTC CGTGTTCGA AGCATTGTA ATGAATTGAG
4951 CATTTAGAAG AGAACAAATT TCTGTTAAG TTTCTTCTAGA TTTTAGATGG
5001 AAAGAATGTA GAAATAAGAG TAGAATGTAG AAATAGGTAT AAAGAATATA
5051 ATAGCTAACC ATTAACAAAGT GTTCCAGAAAT TATCCAGGGAA AGAGAAAAGA
5101 ATTCAAGGC AGTCTGAGA CAAAATTAAG AACCAATTGG AAGTGAAGAC
5151 GCTACATTTT TTTTTCTGG TATGACCTTT CTTTCTCTATA TGTTCCAAT
5201 CTCCTCACTA TGAAATTAGT GAAAAATTAA AGTTAAAAT TAGAGAAAAT
5251 TCACATTAAG TTCTCCTAGG ACTCAGTAGT ATAAGGGTAT AGACTGAGAG
5301 TAGAATGTAG TGTGAGAACAG AGGAGATACA GTATTTAAC ATTACTAATT
5351 CTCTTATACT TGTCTAGTAA TCCTATTCTC TTTTAAAGT CTTCAAGTTAT
5401 TTTCTCTTTA CGCACCTCCT TCTCCCTCTT GTCTCCTCC TTCTACCCCC
5451 ATCTTCTTC CTGTGGAGCC TTCATGAATG GGATTAGTGC TTGTATAAAA
5501 GTGACCTGGA AGACCTTCCT TGCCCCCTCC ACCATGTGAG GACACAGTGA
5551 GAAAACAGTG GTCCATGGAA CCGGAAAGTG GGTCTCACT AGACAGTAA
5601 TCTCCTAGCA CTTCGATCTA GGACTTCCAG TGTCTGGAAC TGCAAGAAA
5651 CAATGTTAT TGTAAAGTA AGCCAGTAGT ATTTTGTCAGA TAGCAGGCCA
5701 GTTGGACTAG GACAATTACC AAGAGCAAGA AGGGAAAGCAG CAAGCTACAA
5751 GAGAGTTCCG TCCTTGGTGT AAATTGACCG TGTATCCTT GTCAAGTTG
5801 AGCCTTACTG GAGCTTTACT TTCTTATCT TAAAATGCAG ATATCTGCC
5851 TGCATCCTGG ACAGAGCTTT TAACAAGTC ATATGTTGCA GAATATGAA
5901 GTTCATGTTA AAAAACCTT TAAAATGTGG TATCCCATT ACTAGCTGGT
5951 GAACTCTTG AGGAACCTCT GTGCCCATGG GTATGAAGTG TATGCTGAAT
6001 GATCACCCAA TGTAGAGGA GTGGGTGGAC TGGTAACCTG ATTTAAGGGC
6051 CATTCTAACT CTACATTCT ATGATTTTT TAATTCTGTC TTTAAGTTT
6101 TACATTTACA ATCACAGAAA AAATAGTCAC ATAGAAGAAT AGTAGCTTAG
6151 CAAATGTTTA TTGCAATTGAG TGGAAATCAGG ATTTCACTCC ATTAAGTAAT
6201 TCCTCTGTTA ACAAAAGAGGG TTCAATTCTAT TTTTATTCTA TTAATATTGC
6251 TTTTTTTTTT TTTTTCTGG AGACAGAACATC TTGCTCTATC ACCAAGGCTG

open 3' end 5' end 3' end 5' end 3' end 5' end 3' end 5' end

6301 GAGTGCAGTG GTGCGATCTC GGCTCACTGC AGCCTCTGCT TCCTGGATT
6351 AAGCGATTCT TGTCCTCAG CCTCCCAAGC AGCTGAGATT ACAGGCACAT
6401 GCCACCACAC CTGGTTAACT TTTGTATTTT CTAGTAGAGA TGGGATTTG
6451 CCATGTTGGT CAGGCTGGTC TTGAATTCTCT GCCTCTAGT GATCTGCCCTG
6501 CCTCTGCCCTC TGAAAGTGCT AAGATTACAG GCATGAGCTA CCATGGCCAG
6551 CCCATTCTC TAATATTTTA ATTGTCAAGAC ATGTTATGGT TTCTGGCACA
6601 ATATTAAGAA GACATGATAT GAAATCACAG GGTGAATTAA AGGGCATCAC
6651 AACAGAAAAGA TTATGGTATA AGAAAAAACAA TGGAATTCCA ACTACATTTC
6701 TGTCAAATGT TCTAAAATAT ATAAAATCTG TATCTTTGT GTTCTCTCCT
6751 GATTTATATT CTAATTGTA TGTTATCCTT CTCTGCAGAA ATAAAGTGT
6801 TGAAAGAATG AAAAATGG AAGAATTCTT TAGTAAGGTA TAAAATACCC
6851 TTTCTATCTT TGTAGCATTCA TAAGCCTTT GTCACCTTC CAAACTCCCA
6901 ACATGCCATA TTCCCTGACT AGGCCACAGC CATGTACATT GATCCCTTTA
6951 TTTTCTCTC TCTGCCTGAG ATTCTCTCA TTCCCCCTTC TCTGCCTGGT
7001 ATATGATTGC CCATTGTTA AGGCCCCAAC TCACCTTTAT AAATCTTCCTA
7051 GCCCACCTTC TTTATCGGTA TTCCAGAAA AACAAAGAA GCTTCCACAA
7101 GACAACATTC TGTAAACAC TGCTTAACCT CTTTGACCC TGCTGAGTT
7151 AAAAATCTTA TCTTTTAAG GATTGAATGG AGTCCACCAA GGTATCTATA
7201 TTTGACAGGA TTTATGAAA CAAAAGGATT TGTTGAGAAA GTTGAGGCC
7251 TAACTCTGAA ACGTGGATCA TAGTGTTCAC TACACATTAA CTGTTTTAGT
7301 GGATGTAATA GTTATTATTA TAGGCTGTGG AACAGAACAA GGGTTCAAAT
7351 GTTTCACCG CTTGCTAGAC TGTGGCCTTG GGCATGTTAT TTAATGCCTG
7401 GAGGCCTCAA ATGTTAACTA GGAATGGTAA GACCTACCC GAACTTAGC
7451 ATAAATAGTA AATTCAATTCA TTTAATGTT TCAAAACAGT GCAAGACATTG
7501 TTTAATGAAC TGGGGATATA GTGGTGAACA ACACTGACAG CGTTCTTCAT
7551 TGTATTCTCA AAACCCCTCCC TATAGTAAGT AGGTCTGTGT GTGTGTGTAG
7601 GTGCATGGGG AATAAAAAAT AATAAGCAAA TAATGAACAG GGTAATTTC
7651 AAAAGCAGAA AGAGCTATTCA AACAAAACCA CCTGCCTTTT ATTAGATGAA
7701 ACTCTCAACT CTATGGTTT TTCTCTCTG TCAATTCTGT TAAATGCTGT
7751 CAGCCTGTT TCCTTATCAC CCTGGCCACAG ACTTCTGTCT TTTCTGCTTG
7801 GTCTGTAGA CTCTAACCCCA AGGCTCATTC TCTGCCTGGC TATCTGCCTT
7851 CTGTGGCTCT TTGCCACTAC CTACATTTTC TGTTGTCAC AGGGAAGGAC
7901 CATTCCCTGT GGACCATAAA ATTCTCTTTTG TGAAAGAATT CATTCTTGAT
7951 TGGGCCACAG CACATCTTGT GAAACAGCAT TAGACATTG CCACTGCTCA
8001 GCAGCTCTGG GGGAAAATGT TTACTGAGAA GCGTACAGTA GTTTTTTTGA
8051 CTAACCATGG TGCAACCTCC TCCCAGAGGG AAACCTATGA GTATTTCAAG
8101 GACATGTGAT GGTCTGTTT TGTCCTGGT ATCTGACATG ATGGGTAGTG
8151 TAGAGCAAGA GCTTACAGAT AATGGCTAAA TTAAATTTC TTTTTGAATT
8201 TTAATATTCA ACTTTTGTAG GTACCCAATC TCCATATTAA GGAAAATAAA
8251 TTACATAAAA AGTGGAGAGT TTTTATTGTG AAACCTGCACC TCCATATTCC
8301 CAGTGGTGCA GGATGAGGGA GCACAGGTGT TGGTCTGGGG AAGCCAGGGC
8351 CCTCTGTGGT TCTGGAGGGT GAGGATTAAG AGGAAGCCTT AGATAGTATT
8401 TATGAGTATC TGCTGACTTC TCTCTGGAC CCAAGATCAC TGAACCTTTG
8451 CCTATTCTGA GATCATCTT CCAATCCAGC CACTAACAGC TGAAGGATAG
8501 GCTTGCCTCT GAGCATTGT AGTGGTTGGA TGAAGATAAA AGATAAAAAAA
8551 CTGTGAGGGG AGGTGTACA GAAGAAAGGG CCCATGTGGG CAGATTTCA
8601 TTCAATTCTC AGTCTTTTATT ACAGCAATTCA TCCAGTGCTG CAACCTTAGA
8651 AAAGGATTCC TACAACACAA TGTAGGTACC CATCAGCAGC AGATTGGATA
8701 AAGAAAATGT GGTACATACA CACCATGGAA TACTATGCAG CCATAAAAAA
8751 GGAGCAAAAT CATGCTCTT GCAGCAATAT GAATGCAGCT GGAAGCCAAT
8801 AACTTAAACG AATTATTGTA GAAACAGAAA AACAAATACT GTGTTCTCAT
8851 TTACAGGGGG AGCTAAACCT TGGGTAATG GGGCATAAAG ATGGGAACAA
8901 TAGACACTAG GGACTCCAAA AGGGGGGAGG GAGGGAGGAG GGCAAGGGCT
8951 GGAAAGCTTC CTACTGGGT ACTCTTCAACCTGGGT ATGGCACGAT
9001 TAGGAGCTCA AACCCCAGTA TCACACAGTA TACCCCTGTA ACAAGCTGAT
9051 GGTGTAACCC CTGAATCTAC AATAAAATTA TTTTATTAA AAAAATCAT
9101 ATAAGGATTG TAAAAAAGAA GGATTCCTAG ACAGGTGCAG CAAACAAATT
9151 TTTTTAAAT GTTGGCAGGC CGCCACCGCC AGTCACTTAT GCTGCAATAG
9201 CCCATGTCCC AACATTCCCA ACCTACTTCT CTCCAAAAGA GAAGCTATA
9251 TTTCAGATGG CCCTGTGCTG GGTTCTCCCT GGAAGTTCT GGGGAAAGGG
9301 GCTTGAGTTG CCCGACTGG ACTCTTCTG GAGTGGGAGC CGGGGCTTCT
9351 GATCAGACGT GAGTGGAGCA GGAACCTCCGC GGTCTCCCAG CGCAGCCCCAG
9401 AGTGCCTGCC CACGCAGGTC CGGGTCTG CGCGCTCGCG CCTTGCCTG

9451 GAAGCCGTTA GGATGAGGCC TCTCCTTCCA GAGCTTAAC CGATGAAGGT
9501 GCATTGTGTT TGGCGCCCT GAGGAGGATG CTGTCCTTAGG CCTCTTCCCA
9551 CTGGACGTGT GTGGTGGGCA GAGATCCCCTG TCGTCGGTCG CACTTCCACC
9601 CCGCTGGGC TCACTCAGGC CGCGGAGCTG CGAGGGAGAC ATCCCTCGATG
9651 GACTCCCTCT ACGGAGATCT CTTTTGGTAC CTGGACTATA ACAAGGATGG
9701 GACCTTGGAC ATTGTTGAGC TTCAGGAAGG CCTGGAGGAT GTAGGGGCCA
9751 TTCAATCTCT AGAGGAAGCG AAGGTGGGTC TCACTGGGC TGTAATCAGA
9801 GAGACGTTGG GGCTGGGAGC CCTGGAGAGG CATTGGGCAG AGAGGGCAAA
9851 ATTACATGT TGTCAGCTT GACCTGGGCC CACTGCAGTG TTCAGGTGGT
9901 TGACCAACGT TACCGTTAT TAAGAATAAC AACACAGCTA ACACATTCT
9951 CAAGTATTT TCTCCGTTT CTCCCTGGCT GTAGTAAAAT CTCCAACCTTC
10001 AGATTGCTCT CAAGATGTT GCTACATACA GCCTTGTCTT AGGAGTCACC
10051 TTGTTCAATG TGCTCACCTG TCATTAGTCA CCCAGAGGG CGTCTAGGCT
10101 AAAGATGCGC CCTCCCCAGT TCAGAGAAC GGAATAATCA CTCTACGTGT
10151 ATTGGGAGT GGGGTGGTGA TTGGAAATT TCTGATGTTA TGTGTTGGTT
10201 TCTGTTCTG GAAGGGGGCA GTGGAAGTGG CTTTTACTCT CGGGTTTCAC
10251 TAGTGTGAG GTTCCTCAT AATATGCCCTT AATTGATAGA CCCTAGTTAT
10301 CAGTACCGAG CTTAGGCTAA CCCTTCTCTT CCCCAGAAGG CTAACCTACA
10351 GGCTCCTCT CAGCATGTT TGCTTCGTAC ATACTCCTAT TGCAGTATTT
10401 CCAAGTCATT TTTCATTGG AATTATTAT TGTATATAAT AATTACTTTA
10451 TAAGTATATT TGCTCTTGG ATGTTTGACC CGGTAGACTG GGAGATCATG
10501 AGCATGTGGA CTATTGAGTT TATTTGGAT AATTGGTACT TCGTGCCAA
10551 AAAACTGTCA GTTGAGTTCT GTCATGTTGA AATTAGTAA AACTCTTCT
10601 ATTAGCCATG TGAACCTTGG GAATATTGAA GCATCCATTG AGTCATGGT
10651 CAGTTCTAGT TTGAGCACAT TCTATATTCC AAGCCCCATA CCCTGGTATC
10701 CTCATCTGTT ATATCAGAGG CCTGGACTGT GTACTTCTG TGGACCAATT
10751 CAGTCCAAAA TGTATTTCT GCAAAGCTTA TCTGGATTTT TAATTCTTAG
10801 AAAAAGCAG TGTTCCTCT TTTAAAGTTA AGTGTCTTG TTCAGGTGCA
10851 GTGGCTCATG CCTGTAATT CAGCACTTTG GGAGGCCAAG GCAGGTGGAT
10901 CACTTGGGGT CAGGAGTTCA AGACCAGCCT GGCAATATG GTAAAACCCC
10951 ATCTCTACTA AAAATGAAA AATTAACCGG GTGTGGTGGT GGGTGTGTGT
11001 AGTCCCAGGA GGCTGAGGCA GGAGAACAC TTGAGCCTGG GAGGCAGAGG
11051 TTGCAGCAAG CTGAGATTGC ATCACTGCAC TCCAACCTGG GTGACAGAGT
11101 GAGACTCCAT CTCAAAAAGA AAAAAAAAAA GTTAAGTGT CTTCATATT
11151 GTTTAAAGAC ACTCTTATAT TTAGATTGTC AAGTGTAAAGT TGTATTTGT
11201 TATTGATAC AAACTAGCCT TTCAAAAGAA ATTCTGGGTT AGCTATCAAG
11251 TCGAACCTTT TGAAACACAT TTCTCCCTA TTGAAACAAA AGGTTTGTAG
11301 AGCTGTCTTG CATTTTGGC AAGGACGCTT TGTGTACCTA GTGGTGAUT
11351 AGGAGGGTTC ACATGTCAA ACCAACGGG GGGGTGTCCC CAGAGAAC
11401 TGCACCAACC ACACAGAAC TTCTGTTCA GAGGAGCACC ATTGTGACTT
11451 TTCCTCAAGT GGCAGTCACA TCGTTAGGAG GTTTGATGT GAGGTCTCTT
11501 CCCACACGTC TCCACCTCCC CAGTAGGAAA ATTGTTTAT ATAGACAAAA
11551 CTCAACTGAT AAAAAAAAAA AAAAGAAAT GATACTTACA TTGTCGTGTT
11601 AAGATACAAA AGCAATAACT TTTTATTGTA AAAATAGTCT GTTTTTGAAC
11651 AATATATTGT TTTGTTTTT CCTGTGAAAG TTGAGAAAAT AAATATACGA
11701 AGAGATAATG GTCAGACCAT AAATAAAAAT AGAACCTTGA CTCAAAATT
11751 ACAGCAGTCT GCCCAGAAA CCAGCCCTT ATCTAAAATA AACAGACCA
11801 GAAACCAGCC TGTATGTCA GACTTATAGG AAGTCAGGTT GCTATCTTA
11851 GAGACAATAC ACAAAAGCTAT GCAATAACTG CTGTAACAGC CCAAATGGT
11901 CAGAATTGTA TTAATAACCG ACAGCCCCC TAATTTTTT CTCACNN
11951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNT
12001 ACCGCTTGCT AGAACTGTGG CCTTGGGTC TGTTATTAA TGCCTGGAGG
12051 CCTCAAATGT TAACTAGGTA ATGGTAAGAC CTACCCAGTA ACTTAGCATA
12101 AATAGTAAAT TCATTCAATT AATGTTTCA AACAGTGCCA GACATTGTTT
12151 AATGAACCTGG GGATATAGTG GTGAACACA CTGACAGCGT TCTTCATTGT
12201 ATTCTCAAAA CCCTCCCTAT AGTAAGTAGG TCTGTGTGTG TGTGTAGGT
12251 CATGGGAAT AAAAAATAAT AAGCAAATAA TGAACAATAA ATTATTTTA
12301 TTTAAAAAAA AAGAAATGAT ACTTACATTG TCGTGTAAAG ATACAAAAGC
12351 AATAACTTTT TATTGTGAAA ATAGTCTGTT TTTGAACAAT ATATTGTTT
12401 GTTTTTCT GTGAAAGTTG AGAAACTAAA TATACGAAGA GATAATGGTC
12451 AGACCATAAA TAAAATAGA ACTTTGACTC AAAATTACA GCAGTCTGCC
12501 CAGAAAACCA GCCCTTATC TAAAATAAC AGACCAGGAA ACCAGCCTGT
12551 TATGTCAGAC TTATAGGAAG TCAGGTTGCT ATCTCTAGAG ACAATACACA

FIGURE 3, page 4 of 42

12601 AAGCTATGCA ATAACGTGCTG TAACAGCCCC AAATGGTCAG AATTGATTA
12651 ATAACCGACA GCCCCCCCTAA TTTTTTTCTT CACTTCAAC TTAGGACGAA
12701 CCAGAGAAAG CTAATATGCA ACCACCTACT AATCAAATAG GGTGCCGCCT
12751 TTCTAATGAA CCCTCCTACA GCTTCCCCAG GCCAGCAGCC CCCAATCAGG
12801 AACGCGCTGA AGCCTCCCT TTTTCTCACT GTAAAGCTTT CCCACTCCTC
12851 TGCGTGGCTT TGAGTCTCTG TCAATACACA AGTGAGGGTG TCTGACTCCC
12901 TTGCTATAGC AAACCTGGGC CAAGTAGATT TTACTTTCT CATTGATTG
12951 GTCTTTTATT TCTAGAAGGA ACATACAAGA AAATTAAAG GGGATCCAT
13001 TCCTAATCTT TCATATTATA GTAGTCCCCCT TTTATCTGCA GGGCATATTT
13051 TCCAAGACCC CCACCTGAATA CCTGAAACTG TGGGTAAATAT TGAACCCTAT
13101 ATATACTCTC TCTATATATA CATATATATA TATATTTTT AATTTTTTT
13151 TACTTTATCT TTAATTAGCT TAGTCTTTT TTTTTTTTG TGAGATGGAG
13201 TCTCACTCTG TCACCCAGGC TGAGTGCAGG GGTGCAGTCT TGTTCACTG
13251 CAACCTCTGT CTACCGGGTT CAAGCAATT CTTGTGCCTC AACCTCCGGA
13301 GTAGCTGGGA CTACAGGGCT GTGCCACAC TTCTGGCTA ATTGTTTAA
13351 ATTTTAGTAG AAACGGGATT TCACCAAGTT GGCCAGACTG GTCTCGTACT
13401 TCTGACCTCA AGTGATCCGC CCACCTTGGC CTCCCCAAACT GCTGGGATTA
13451 CAGGCGTGAG CCACCATGCG CCCAGCCATA GACTATATAT TTTTGTCTG
13501 ATAACGGTT CAGCTACTAA GTGACTAACAA GGCAAGTAGC ATCTATAGTG
13551 TGGATATGCT GGACAAAAGG ACATTCACCT CCTGGGCAGG ATGGCACAGA
13601 ATGTTGAGAG ATTTTATCAT GCTACTCAGA ATGGTGTGCA ATTTAAAAC
13651 TATGAGTTGT TTGTTCTGG AGTTTCCAT TTAATAGTTC AGACCATGG
13701 TTGACCGCAG GTAACGTAAA CTGTGGAGAG TGAAACTGTG GATAAGGGAG
13751 GACTATTGTA TTGTTAACGTC AGACTCATTA GGCAATCATA ACTCTTGATT
13801 TGCCATCAGA AATGCTGCAG AAATATGGGT TAAAAAAAC TGTTCAAAAA
13851 TAGGGTCAGG GATGTCCTTT AACTTGTAC TTCCAAAATG TTAGTAAAAA
13901 CTGTGGCCCC AAAGAGTGA AGGAACAAAT GACTAAGAGA AAATCTGTT
13951 TTCAGGATGA CAGATTTAAAGAAGCAAC TTGCTGAAAC ACTGAAAATC
14001 TCTCCACTTG TAAGATAACA CAAAACGGC TAAAACGGT TGGAAATGAAT
14051 ATGGCCAACCT CAAGCTGCA CAGAACTAAC TTGGTGTATGT TACAGCCAA
14101 ATTTCCACCA CATATTTAT ACTAACTCCC CCCGGATTTT CACACATGAT
14151 CTGTGAGGTA GCATGAAGAG GTAACTATGC ATGCCCTAAGG ACTTGGGAGA
14201 CCTCCCCATT TCCTTCCACC AATCACCCAC TAATCCCAGA ATCCGCCCCC
14251 AAACCTTTTC TAATAACTAC CTTAAAGCCA GCATAGGGAG ACAGATTGA
14301 GCTGGACTCC TGTCTTCTTG TGGGTACCT TGCAATAAAAA AGCTTTTCTT
14351 TTCTCAACAC CTGGTATTAT AGTATTGACT TCTAGTTCAT CGGGCAGCAA
14401 GCCCCTTTG GTCGGTGACT ATTCTTGTTC GCTGATATT CCATTGGCCA
14451 AAATATAAAC CTCTTAGATG AAACCTTCAGT ACGTAAATGG CGCCACAGAA
14501 TGCTGTGACA TTTTCTCTT GGATTATAGC AGGTTACTTT ACTGAATACC
14551 GTAGGCAGTT ATAACACACT AAGTATTGTT GTATCTAAC ATAGAAAAGA
14601 TACAGTAAAAA ATATGGTAAT TTTTTCAAC TTTTAGTTGA GATTTGGAGG
14651 GTATGTGCAC ATTGTTACA AGGGTATATT GCATGATGCT GAGGTTGGGG
14701 GTACATTGA ACCCTGTAC CCAGGTAGTG AGCATAGTAC CCAATCGATA
14751 ATTTTCAAC CCTTGTCAC TCCCTCCCCG TTCTGTAGT CCCCAGTTTC
14801 TGCTTTCCC ATCTTTATAT CCGTGTGCAC CCCATGTTT GCTCCCATGT
14851 GTATGTGAGA ACTTGTGGTG TTTGGTTTC TATTCTGCG TTGATTGCG
14901 TAGGATAATG GCCTTCAGCT GCATCCATGT TGCTGCAGAG GACGTGATT
14951 TATTCTTCTT TATGGCTGTG TAGTATTCCA TGGTAAAAA TATAGTACTA
15001 TAACCTTAAC AAATCACTGT CATATATATG GTCTATCATT GACTGAAATG
15051 TATACAGTGC ATGATATATA TATATATATA TCTATAATGT CTTATCCATT
15101 TCGTGTATTA TGAGATTGTTA TTGCTAATAT TTTATACAGG AGTTTGCAT
15151 CTTTTCACCT AGTTGACATT GCTTGTAAATT TTCTTTTTT TGTGATGTCC
15201 CTGTTAGGTT TTAGAATCAA GTGTATACCC GCCTCATAAA ATGGGTTGGA
15251 AAATGGTCCC ACCCTTCTG TTCTCTGAA AATTGGTGT TTTTTCTTAA
15301 AGTTGGTAG ACATTATTGT TAAACCATG GGGTCCTCGA TTTTTCTTCA
15351 TGGAATGTT TCAAATTAC ACTTTAATT TCTTAAAT CTGAGTATAG
15401 GGCTATCAGA CTTTCTGCTG TCTTATGTCA GTTTTAATA AGTTGTTTTT
15451 GTAGGCAGTT TTGATCTCAC TTTCATATT TTGATATAAA GCTTTTCATA
15501 ATATCATTAA TGTCTATAGT GTCTAGTAGT TTCCATCTT ACTTTCTGAC
15551 ATTGGTTATT TGCCAGTTTT AGGAGTTAT CAATTTATT AGTCTTTCA
15601 AAGAACCATC TTTGGCTTT GTTAATCCTC CCAATGGTGT GTTTCTTTC
15651 TCATTACTTT TTGCTCTTTA TTCTCTCAA CTTCTTTTTT GCTTAATTTT
15701 AAAATAATT CTTGAGATTG AGATAAGCCT CAATGATGGG TCACCGATT

15751 CCAGTCTTC TTCTTTCTA ATTATGCATT TAAACCAGA AATCTTCTC
15801 TAAGTGTAGC TTTAGTTGCA GCTCACAAAGT TTCAGATCTG TCTCTCAGTC
15851 TGGAGGTTGG AGATCTGACC ATGACCATGA AACCATCCAG TCACAATGTG
15901 GCATTATTTT TTTAATTNTT TTTTTTTTT TTGAGATAGA GTTTCACTCT
15951 TATTGCCAG GCTGGTGTGC AATGGTGCAG TCTCGGCTCA CAGCAACCTC
16001 CACCTCCCAG GTTCAAGCGA TTCTTTGCC TCAGCCTCCC AAGTAGCTGG
16051 GATTACAGGC ATGCGGCCACC ATGCCAACT AATTTTGAT TTTTAGTAGA
16101 GATGGGGGTT CTCCATGTTG GTCAGGGTGG TCTTGAACCTC CCGACCTCAG
16151 GTGATCCGCC CACCTCAGCC TCCCAAAGTG CTGGGATTAT AGGAATGAGC
16201 CACTGTGCCG GGCCAACAT GGCAATTAT ACCCAGAAGA GCATGACCAT
16251 GAGAACAGTA GAATTGTAA GCTTGAGTG GGTGACTATG AGTGTATAA
16301 TAGGTAGATA GGTATATTT TGGGTGGTGG TAGGAGAGGG CTTACAGTTT
16351 GCTATGACAG CTTTTTATAT GGATCATCTC TAGAAAAGA TTATTTAATT
16401 TTTGAATCA AAGGGGAAAA CACTAGTTA GGCTTCTTC TTTCTTCTT
16451 TTTTAGAGC AGGGTCTTGC TCTGTACCA GGTTAGAATG CAGTGGTGC
16501 ATATTGCTCA CTGTAACCTC AAATTCTGG GCTCAAGTGA TCCTCCTACC
16551 TCAGCCTCCA AGTAGCTAGT ATTTACAGGC ATGCACCAAC ACATCTGGCT
16601 AATTAAATAA ATTTTTATG GAGATGAGGT CTCACATATGT TGTCCAGTCT
16651 GGTCTGAAT CCTGACCTCA AGTGATCTC CCCCATCAGC CTCCCCAAGT
16701 GCTCAATAT TTTAAATCCT GTGGTAGGTC AAGTGGTTGT CTTCTATCTT
16751 GGGGTTATA AAGTACATGT CAAGAAATT AGGGTATGGT TAGATTAGCT
16801 TTAAAAATGT CATGTTTAT AAAATCAAT GCATCATTTC TCTGATTGAA
16851 AATTAAACAC AAGACTCAGA ATCTTTTGC AGTAGTGGAA TTACTTTTAT
16901 TATAGATCTT TCGGATAATG AATGATGATA CATCTGGCCA AAAATAGTA
16951 CTATAGTCTT TTAGAAAAAC AGCTAATCTG CTTGAAATAT GTGTAGAAAT
17001 AATTAGTGC ATCAGCCCAT ATTGGCAATA ACTTCTCTCT AATTTTTTTT
17051 TATAGAAAAT TTTTACTACT GGAGATGTCA ACAAAAGATGG GAAGCTGGAT
17101 TTTGAAGAAT TTATGAAGTA CCTTAAAGAC CATGAGAAGA AAATGAAATT
17151 GGCATTTAAG AGTTTAGACA AAAATAATGA TGGTGTGTCT TTCTTTGTA
17201 TTTATCACCA GCTATGAAGA AGCATTATC ATGCTTCAA GAGTCTAAAA
17251 GGATGTTAT TTAATCTCTC TGGTTTAAAGA TGATAATTAT TATTTGTGTT
17301 AATACTTTTT TTTAGTAATG TGATTTTTAT GTAGAGTTA TATTATTTAG
17351 TGAAGAAAAC TTATAGATAG CTTTTCTTT TCATTACTTT GAAATGTAAT
17401 GAATTACATT TCTGAATTAA AAACCTGGGG CAGGGCCTGT TGTAAATGTT
17451 AACTATGGAA CATTATGCTG ATTTGAGTTA AACCTGTAGG TTAAAAAATAA
17501 TAATTATATT TTCTTGTCTC CTGGGTAAAA TGAGATTCT TTTTATTGTT
17551 ATAGAAGAAT GACAGTTGTG TCATCTAAAA TTTAAAAAAAC TTTCAGATTA
17601 TCTTGATCT GTTAGTTTT TTGGAAGAAT TAATTAGAG AAGATATCTC
17651 TGATCCTGGA AATTAGGGAA AAAAGCATA TAAACGTTA AGTGTGTACC
17701 TTCTGGTTAA GATTATGACT TCTATATTTC GATTAATAGG TTGGAGTTG
17751 TCTTAATCTG TTTTCTGTTG CTGTAATGGA GTACCCACAGA CTGGGTAAATT
17801 TATGAAGAAA TGAAATTAT TCTTATAGT TCTGGAGGCT GGGAAAGTCA
17851 AAGTTGAGCC GAATCTGGTG AGGGCCTCTT ACTATGTCAT AACATGCTAG
17901 CAGGCATCAC AGAGCAAATG CACTACCTCA GATCTCTCTT CCTCTTCTTA
17951 AAAAGCCACT AGTCCCCTCA TGGGGCCCT ACTCTGAAGA CCTTATCTAA
18001 TTCTAATTGG AAATAGGGTC TTGAAGCCT CATCACTAGA GGTAACCTTT
18051 AACAGGAAGA GAGAATTAT AAAAATTATA ATGCAGCACC AAATCCCTCC
18101 CTACTTGTGA ATAGTCAGG TCATTTCATT TACAGACTTG TTATTAAGA
18151 AACAGGTTAA ACAAAATAGAT TGAGAGGAA TGTGGTTCAT GTCTGAGATC
18201 AGCAAACCTT TTTGTCCAGA AGTCCAGATA ATAAATATTT TAGCTTTGTG
18251 GGTCTGTGG TCTCAGTTGT AGCTACTTGT CTCTGCTGCT GTACCTCAAA
18301 AGCAGGCCATG GATAATATGT AAATGAATGG GGATGACTGA TTTCCAATAA
18351 AAAACTTATT TACAAAGATA GTTAATACAC CTTATTTGGC TTGAGGGTTA
18401 TAGTTGCCA TCCCTGTGATT TACAATGAAT ATTAAGTTT AATTCAAAGC
18451 AAGTTCCCTC AAACAAACAA ACTAAACTCT AGATGATTTT GAAGATTATT
18501 CACATCTGTG ACTCTCAGCC AGGAAGAGCT GAGTTGGGT TGGAAAGTAG
18551 TACTATTGGA ACATTTGTG CCCATAAGCC TTACAATATA TGCCCCCTAAG
18601 TCTAGCCTTA GTCCAGTCTT CTAGCAAAAC TCAGTTTCTT TTCTTCTCTG
18651 CAAACTTCA TTCCAACATC GACCCCTCTG AGTTCAGATT GTCTTGCAGG
18701 TCAGATTGTC TGTGTGCTGC TATGGTAGGC AGTAGCTGAG AGATGGAGCT
18751 ACCTTAAGAT CAATTGCCAG ATAATCAGAG GTCAATTATC CCAGTGCATA
18801 AGTAGTGTAC ATATCAATTG TTCATTATAT AAAATTCTAA ATGAACCAGA
18851 GGCAATAATT AAAGATGAAA TTTTGATGGT ATATTGTAG GAAATCTACA

FIGURE 3, page 6 of 42

18901 CAATGTTCC CTAATTCCC ATGTTTGTG ATTTTAAAC AATGTGGCAT
18951 TATTGGTC CATTTTATT TTCTAGACTT CCTTAATGCA AAACATATA
19001 AGTTGATCCT CATTATTTGG GGATTCTGTA TTTGCAAATT TGCTACTCA
19051 ATAAAAATTAA TCCCCAAAGT AACCCCAAAA TATATACTCA CAGTACTTTC
19101 CCAGGCATTG ATGGACATGC ACAGAGCAGT GAAAAACTTG AGTGCTCAG
19151 CATGTACATT CCTAGCTAGT AGAATAAGGC AATACTCTGC CTCTTGTGTT
19201 CAGCTCTCAT ACTATTAAC AGCAAGTATC CCTTCAGG TCTATTTGT
19251 GCCAGTTTT GCATTTGTG ATTTTGTG GTAATTCCCT TTTTAAATG
19301 TTCCCCAAAG GTAGTGCTGA AGTGCTGTCT AGTGTCCCTA AGTGCAGAA
19351 AGCCATAGCA TGCCATTGG AGAAAATATA TGCCTGGAT AAGCTTGCC
19401 CCAAATTCAA TGTTAGTGA TCAACAGCAC ACATTAATG AGGTGCCTTC
19451 AAACAGAAAC AGACATAAGA CATGGTTATG TATTAATCAG TTGATGAAAG
19501 TGTTGTAATC AGAGGCTCAC AGGAACCTAA CCCTGTTTT CCTGTAGGAA
19551 CAATGGTTG GTATTTGCTA ATTCACTGTT TGCAATGAAT ATAGAACTTT
19601 ATGGAAGATG ATTGCTGTGA ATAATGAGAA TTAACCATAT CTCTTAAAGA
19651 GTGCATTCT AAAGGAGAA ATTCAAGGG GTATTTGCT AATTTCTTTA
19701 CTAACAGATG CTGCCCTCA CTGCTCTAC ATGGTCCAGA TTCTCATGCT
19751 GCTCCTCCC TCTCCCCAGG AGGATTCTCT CAGAACCTG TCATCTCCTC
19801 CAGGGTCTT TCTCCAAGAA AGTCTATCCT TTCACCACTA ACAGTAATT
19851 TGGTCTCCT CTTTTCTGG AGAAGTCAGC TGTTTATGCT GCTTCAGCAC
19901 CAGACCTCT CTTACTTTGT TTTGTTCAT TCTTTTCTAT GTACAGTAGT
19951 CTTAGGATTTC TCATGAGCCT GTGAGCTGCT AGAAGGAAAT ACAGCAGTGC
20001 TTACATTAT TGCTTCTATT TTATTTCTA TTTCTCTTC CTGCTCTCTG
20051 ATTGTTCTCC TTCTGTCCAC AAACATGTC TAATTCCTCT AGTATTAAGA
20101 ATTTCTGTC TTTGTTGTT CTTTATCCT TGCTCCCTA TTTTTACTGC
20151 CAGATTTTA TTTTATTTA TTATTTTTG AGATGGAGTC TCACTCTGTC
20201 ACCCAGGCTG GGGTGCAGTG GCGCGATCTC AGCTCACTGC AACCTCCGCC
20251 TCCCAGCTTC AAGCAATTTC CCTCTTTAG CCTCCAAGT AGCTGGGATT
20301 ATGGGCACCT GCCACCATGC CTGGCTGATT TTTCTATTT TAGTAGAGAC
20351 GGGGTTTCAC CATGTTGGCC ACACTGCTCT CTAACGCTG ACCTCAGGTG
20401 AACCAACCGC CTCAGCCTCC AAAAGTGTG GGATTGCAGG TGTGAGTCAC
20451 TGTGCCTGGC CTTTACTGC CAGATTTTA AAAGAATAGT CTGTGCTTTA
20501 GCTCTATTC CTCATTTACT ACITCTCTT AACTCAGTC TATATGATGTT
20551 TTGCTATAGT AAATGTCAG TAATTTATTA AAAATGAGA AATAGGTACT
20601 TTAAATGTA ATAGATCCTA CTTTAATTGA ATTTATCTTG GAGTTAGAAT
20651 ATCTTGATTG GGATTTTGT TCTGCTACTT CTAAATTACA TTACTTGGTA
20701 AGGCCACTTG TGAAGTCAGT CTCTTGGAG GAATATTATT TATCTATAAG
20751 GCTGTTACAA TTACTGAATT TTAAAAAATG TGTATTTATT TTTTAATGTA
20801 TTGTTACAT TTCTAGTATT GATGTTGGGA TAGCCATTAA AGCAAGTC
20851 TAACTCACCT ACATGCATAA TTTTGCCTTA ATCAGTTAA AGCTTTCTCT
20901 TAAATGAGAG ATTTGAAATT CATAATTCT GTGGTTCTTA TCAGTTCTGA
20951 GTTTTATTTT TTGCCCCCTTT TATTTTTTA AAGGAAAAAT TGAGGCTTC
21001 GAAATTGTCC AGTCTCTCCA GACACTGGT CTGACTATTT CTGAACAACA
21051 AGCAGAGTTG ATTCTTCAAA GGTAAGCTCT TCATGTTGGT CAACAATTGA
21101 CTTTCACCTT AATATCCTGC ATTAGAACTC TGTGTTGTA AGTGTGGCTT
21151 TAAAACACCT CCCTAGTCCT CATTATGTAT ATCCAAGATC TTTTGTCTT
21201 TTTTCTCCC ATTCAATTG TATGTGTACA TTTATCTAAA GTGTAAGAAT
21251 GGGAAAGTGT AGTCTCAGACT GGACTCTTTC TTTCAAGGCC TCAAAGGATA
21301 GTGGAATGGC AGGAAGTAAG GTTTAACTC CATAGATGAG GAGCTGAAGA
21351 GTTTGGTGT TGCTTTCT CCATTTGATT TCTAATGTGA CAGTAAA
21401 CATTGATTCA AACTAAGAAG ACTAGCAGAT TCATCACATT ATTTAACCTA
21451 GATGTGACTG GAAAAAAGGG AAATTACTAA GCTCTCCAAG CTAACAAAGA
21501 AATACCTGTT TAAACTTCA GAAAACAGAA ATGCAAATTG GAACCTTATT
21551 GTCTGGGCA ATCAGTTGTA CTATTTAAGT CAGACTTTA TACTCTTAAT
21601 GTTTGTTTC ATGGGATAGA GCAGTAATCT CTGCAGGCC GGTGCTCTCA
21651 AATACTCTGT TGCTATAAAC ACAGGGCAGG AACTGATTTT TTATGATAAC
21701 GTAAAACAGA AAAGGACAAT TATATTGTAT TAATATTGTT GTGAATATT
21751 TCAGCTCTCA CATTGTCTAA AAATCTTCT AAATGGCTTT GTTATTGAA
21801 TTATCTCATT TTATATCTGT GCCAACAGCA TTTTCATCCT TTCTCTTCAT
21851 AATTCTTTT ACAAACAGCT GCTCAAGAGG AAGGCTCAA GTCTCAAGGC
21901 TGAGCACGTA ATGACTTTG TTAGTACTAG ATGAGAAGGG CTTTCCTGAG
21951 GAAATGAAAA CCTAAAACAT GAAAAGAAGA TAAACAGAAAT TTGGACAGTG
22001 AGATATAGAG CATATAATAT TCTGCTCTA AAGTAATATT CTTCTAGGAA

22051 AGTGAGGGCG TTTCCCTGGC TGTTAGGCCA GAAATCATAT TCCTATATTT
22101 TCTTTGATAG CTTTAGGAAT AATGCAAATT CTAAGCCAA GCCTCAGAAT
22151 AGACTAAGAA GTATTAGCTT AGCTGCCATG ACAAAATACC ATAGGCTGGA
22201 TGCATTAAC AATGGAAATT TAGTTTTCA CAGGTCTGGG AGCTGGGAAG
22251 TTTAAGATGA GAGTGCAGC ATGGTTGGGT TGTAGTGAGG GCTCTCTTC
22301 TGGCTTGCAAG ATAGACCCCT TCTCACTGTA TTGTCATATG GCAGAGAG
22351 AGAGAGAGAG AGAGAGAGAG AGAGAGAGGG GATCTTCTC TTGCTTCTA
22401 TTATAAGGCC ATAGTCCTGT TGGATCAGGG TTCCATTCTT ATGACTTTAT
22451 TTGACTTTAC CCCCTAAGA TGCTATCTCC AGATATAATC ACACGGTGGG
22501 TTAGGGCCTC AACATTGGA TTGGGGAGGG ACACAGCTCA GTCCATAGCA
22551 AAGGATAATG CAGAGGGTTG GATATTTAAA AGTAGCTACA CAATTTTAA
22601 TATAATATT TTATGGTAAC TTTTTTTTT TTTTGAGATG GAGTCTAGCT
22651 CTGTTGCCCA GGCTGGAGCG CAATGGTGC GATCTAGCTC ACTGCAACCT
22701 CCGCCTCCCCA GGTCAAGCA ATTCTCCTGC CTCAGCCTCC TGAGTAGTTG
22751 GGACTATAGG CACCGGCCAC CACGCCCTGGC TATTTTTTT TTATTTTAC
22801 TAGAGACGGG TTTGCACCAT ATTGGTCAGG CTTGCTCTCGA ACTCCTGACA
22851 TCAGGTGATC CACCCATCTT GGCCCTCCAA AGTGCCTGGG TTACAGAACT
22901 GAGCCACCGC GCCTAGCCAG CAGCTTACT GAGATGTAAT TCACATGCCA
22951 TAAATTCACT TTCTAAAGT ATACAATTCA GTGACTTAAA ACATTTATTT
23001 ATTTTAAAT TGACAGAATT ACATGTATT ATCATGTACA ACATGATGTT
23051 TTGAAGTATA TGACATTGT GGAGTGACTA AGTCTAGCTA ATTAACATGA
23101 TACATCTCAT ACTTAATGAT TTCTGTGGTG AGAACACTTT ACATCCATTC
23151 TCTTAGTATT TTCTAAAGAT ATAATATATT ATTATTAATT GTAGTCTTCA
23201 TGTTGTATAG TGGAGCTCTT GAACTTATTC CTCATGTCAA GCTGAAATTG
23251 TGTGTCCTT AACACAAACC ATACCCGACT CCCAAAGTAT TCTGCTCTCT
23301 GCTTCTATGA GATTAACCTT TTCTGATTCC ACATGAGTGA GATCATGCAG
23351 TATTTATTTG TCTTACCTG GCTTATTCA TTCATATTGT TACAGATAAC
23401 AGGATTTCCT TCTTTTTTA ATGGCGAAT AGTTTCTAT TGTATATGTA
23451 TAGCACATT TCTCTCTCA TGCAATTGGT GACACTTAGG TTGATTCCGT
23501 ATCTGGCTA TCGTAAGATAG TGCTATAATG AACATGGAA TGCACATGGC
23551 TCTTGACAT ATTGATTCA TTTTATATAT GTGTATATAT ATATGTATAC
23601 ACACACATAC ATACAGTGGT GGGATTGCAG GATCATATGG TAGTTCTATA
23651 TTTAATTTTT AAAGGAACCTC CATACTGTT TCCATAATGG CTGTATTAGT
23701 TTAACCTCTC ACCAACACAGGG TGCAAAAGTT CCCTTTCTC TACATACTTG
23751 CCAACACTTG TTATCTTTG TCTCTTTGGT AATAGTCATT CTAAGTGTAG
23801 TATGAGGTGA TATCTCATTG TGGCTTTAT TTGCAATTCT GTGGAATTAA
23851 GTGATATCGA GCTTTTTTTT TTTTTGTAC TTTGCCATT TGTATGTCTT
23901 TGAAAAATGT CTATTGGGGT TTTTTGGTT TTTATTTGAG GTTTTNNNNN
23951 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
24001 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
24051 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
24101 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
24151 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
24201 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
24251 NNNNNNNNCCG GGGTCCCGT CATTCTCCT GCCTCAGCCT CCCCAGAGTA
24301 GCTGGACTA CCAGGGCACC CGCCCACAC GGCCCGGGCT AATTTTTGT
24351 ATGTTGAGTA GAGACGGGGT TTCACTGTGT TAGCCAGGAT GGTCTGATC
24401 TCCTGGCCTC GTGATCTGCC CGCCTCGGCC TCCCAGAGTG CTAGGATTAC
24451 AGCGTGAGC CACCGCGCCT GGCTGATT CTAGTTTTT ATTATTGTGG
24501 TCGGAAAAGA AACTTGATAT GATTCATTC TGCTAAATT TGTTAAGACT
24551 TGTTTGTTGG CCTAACATAT GATATCCCT GGTGCATGTT CCATGTGCAG
24601 TTGAGAAGAA TGTGTATTCT CTTGCCATTA GGTGAAATGT TTTATGTCTG
24651 ATCTGTCCAT TTGTTCTAGA GTATAGTTA AGTCTGATGT TTCTTACTGA
24701 TTTTCTGTT AGATGATTTG TCTATTGCTG AAGGTAGGGT GTTGAAGTCC
24751 CCTACTATTG CTGTATTGCA GTCTCTCTC CCTTCAGAC GTATTAATGG
24801 TTTTATTCTT ATTTATTG TTGTTGTTGT TGTGTTGTT GTGTTTTTG
24851 AGACGGAGTC TCACTCTGTC ACCAGGCTGG AGTGCAGTGG CAGGGTCTCG
24901 GCTCACTGCA GCCCCCGTCT CACGGTCAA GCGATTCTCC TGCCTCAGCC
24951 TCCCGAGTCG CTGGGACTAC AGGCGCATGC CACCAAGGCC AGCTAATTT
25001 TGTATTTTA GTAAAGACGG GGTTTCACCA TGTTGGCCAG GATGGTCTTG
25051 ATCTCTTGAC TTCAATGATCC ACCCGCCTTG GCCTCCAAA GTGCTGGGAT
25101 TACAGGTGTG AGCCACCACCC CCGGCAAT GTTTGGTATT TATCTTCTAGG
25151 TGCTCTGATG TTGGGTTCAT ATATATTAT AAAAACAAAT AGCTACATAA

25201 CTTATTAAGG GATATGCAAT ATAAAATATA TAAATTGTGA CACTGAAAAT
25251 TTAAAATGGG AGGAGTGGAG TAAAAGTACC TTCATATAAC TTACTATTAT
25301 ATCCTCTTAT TGAATTGACC CTTTTATCAT TATATAGGAA CTTTGTTCT
25351 CCTTTACAAC TTCTGACTTA AAGTTTGT TTATATGATAT AAGTAAAGTT
25401 ACTCCTGCTC TCCCTGGTT TCTGTTCCA TGGAATATCT TTTCCATTG
25451 CTTCACCATC AGCTGTGTG TATTTTACA GATGAAATGA GTCTGTATG
25501 GGCAGCATAT AGTGGATCT AGTTTTTA ATCCACTCAG ACACTGTGTT
25551 TTTTGATTGG ATAATTAAAT CCATCATGT TCAAGGTAAT TATTGATAAG
25601 TAAGGACTTT GTACTACCAT TTGCTTATT GTTTCATGGT TCTTTTATAG
25651 ATCCTTATT CTTTCTTCC TCTCTTGTG TCTTTTTTT GTGGTTAAGT
25701 GATTTCTCT AGTGGTATGT TTTGATTCT TGCTTTTAT TTTTTGTGTA
25751 TCTCCTATTG GTTTGGTT TGTTGTTACC AAGAGGTTAC AAAAACATC
25801 TTAAGAGTTA TAATAGTTA TTTTAACTTG ATAACCTTAAT TTTTATTGCA
25851 AAAACCCCCC AAAACAAAAA AATCTACACT TTTACTTAAT CCCCTGAAAT
25901 TTTGAATTTT TGATGTACA GTTTACCTCT TTTCATATTG TGATCCCTT
25951 AAATTATTGT AGCTATTATT ACTTTAAATA GTTTCTCTT TCCTACTACA
26001 GATGTAAGTG ATTGCTACAT CATCATTACA GTATTATTTT GAATTTACCT
26051 GTGTACTTTT TTTTATCAGC CAGTTTATAA CTTTCAGATG TTTTTGTGTT
26101 ACTCATTAGC ATCTTTTCTT TTCAGCTTGA GGAGCTCCTT TTACGTTCT
26151 TATAAAATAG GTCGGGTCA GATTATCTCC CTCAGCTATT GTTGTCTGG
26201 GAAAGTATCT CTCCCTCATT TCTGAAGGAC ACTTTGCTGG GTACATTACC
26251 CTTGGTTGGT ATTTTCTCC TTGAACGCTT TAAATATATC ATCCCTTCT
26301 CTCTGACCT GTTAGGTCTC TGCTGACAG TCTGTTCCA ACCATATTGG
26351 GACTGTCTTA TATGTTATTT GCTTCTTATC TTTTGCTGTT TTCAGGATCC
26401 TCTCATGTC TTTGATTTT GATAGTTGA TTGTAATATG TCTTGGGTA
26451 GTCTGTTTG GATTGAATCT GATTAGAGAC CTTGGACTTT TCCTGCATGT
26501 AGATATTAC CTCTTCTCC AGGTTGGAA AATTTCTGT TACTGTTCT
26551 TTAATTAAAGC TTTTACCCC TTTTATCTTC CTTTCTCCT TCTTCAACTC
26601 CTGTGACTCA AAACTTTGCT CTTTGATGC TGTTCCATAA ATCTTGTAAAG
26651 CTTTCTTCAT TCATTTTCAT TCTTTTTCT CCTCTGTGTA TTTTCAAATA
26701 ACCTGTCTT GAGTTCATAG TTTCTTCTT CTTCTTGATC ACTTCTGCAG
26751 TTGATGCTCC CATAATTGCA TTTAATTGGT TTCAATTGTAT TTTTCAGGCC
26801 CATGATTCT GTTTGATTTT TTCTTTTATT ATTTCATCTC TTTATTACCT
26851 TTCTCTTGT GGTCACTCGT TATTTTCTTA ATTTCAATTGA ATTGTTCTT
26901 TGTATTCT TGAAGTTGC TGAGCTTCT TGAAATTCTA TGTCAGTTCA
26951 TACATCTCTG TTTCTTTAGG GATGGTCGCT GGTACTTTAT TTTGTTCTT
27001 TAGTGGTGTCA ATTGTTCTC GATTGTTGTT GATGTTGTG GCCTTGTGTT
27051 TACATCTGTG CATTGAAAGA AGTAGGCCT TATTCTGACTC TTTGCAGACT
27101 GGCTTGTCT GAGAATGCC TTCAACAGTC AGCCTGTCTA GAGATTCTT
27151 AATATTAAAT TAAATATTCT TAATATTGG AAGAACTTCC AAATTGTTTC
27201 TAAAGTGGCT GCACCATTTT ATAATCCAG CAGCAATGAA TGAAGGTTTC
27251 AGTTCTCCA TAGCTATATG AATACTCATT ACTGCTGTC TTTTCATTT
27301 TTGATTTTA TTTTTTTTTT GAGAAAGGGT CTTGCTCTGT CATCCCATCT
27351 GGAGTCAAT GGACAATCA TGGCTCATTG CAGCCTCAAC TTCCCTGGCT
27401 CAATTGATCC TCTCACCTCC TGAGTACCTG GGACTACAGG CATTGTACCA
27451 CAATGCCCTGG CTAATTTTA TATTTTTGT AGAGATGTGG TTTTGCCATG
27501 TTGCCTGGTG TATTAGTCCA TTCTCATGCT GCTATAAAGA ACTGCCTGAG
27551 ACTGGGTAAT TTATAAAGGA AAGAGGTTA ATTGACTCAC TTTTGCTTGG
27601 CTGAGGAGGC CTCAGGAAAC TTACAATCAT GGTGGAAGGG GAAGCAAACA
27651 CGTCCTCTT CACATGATGG CAGGAAGAGC AGTGCCTAGC AAAGAGGGAA
27701 AAAACCCCTT ATAAAATAAT CAGATCTCAT GAGAAGTTAC TCACTATCAT
27751 GAGAACATCA GAATGAGGGT AGCCTCTCC ATGATTCAAT TACCTCCCAC
27801 TGGGTCCTCTC ACAGTGACATG TGGGGATTAT TGGAACTATA ATTCAAATG
27851 AGATTGGGT GAGGACACAG CCAAACCATA TCATTTTGC CCTGGTCCCT
27901 CCCAAATCCC ATGTTCTCAC ATTGCAAAAC ACAATAATGC CTTTCCAGCA
27951 GTCCCCCAGC GTCTTAACTC ATTCCAGCGT TAACCTAAA GTCCAAGGTT
28001 TCATCAGAGA CAAGGCAAGT CCCTCTGCC TATAAGCCTG TAAAATCAA
28051 AGCAAGGTAG TTATTATACT TCCTAGATAC AATGAGGGTA CAGGCATTGA
28101 TTAAATATAC TTGTTCCAAA TGGGAGAAAT TGGCCAAAAT GAAGGGGCTA
28151 CAGGCCCAA GTAAGTCCGA AATCTAGTGG AATAGTCAAA TCTTAAAGCT
28201 CCAAATGAT CTCCCTTGAC TCCACATCAC ACATCCAGCT CATGCTAATG
28251 CAAGAAGTGG GCTCCCATGG CCTTGGGCT CTCGACTCCT GTGGCTTTTC
28301 AGGGTACAGA CCCCTTCTG GCTCTTCA CAGGCTGGCG TTGAGTGTCT

28351 GTGGCTTTTC CAGGTGCATG GTGCAAGCTG TCGGTGGATC TACTATTCTG
28401 GGTACTGGAG GATGGTGGCC CTCTTTAC AGCTCCACTA GGCACTGCTC
28451 CAGTGGGGAC TCTGTGTGAA GGCTCCAACC CCACATTCC CTTCTGCACT
28501 GCCCTAGCGG AGGTTCTCCT CAAGGGCTCC ACCCCCTGCAG CAAACTTCTG
28551 TCTGGACATC CAGGCATTTC CATACATCCT CTGAATCTA GGCAAGGAGAT
28601 CTCAAACCTT AATTCTTATC TTCTGTGTAC CGCGAGACTC AACACCTTGT
28651 GGAAGCTGCC AGGGCTGGG GCTTGCACCT TCTGAAGCCA TGGCCTGAGC
28701 TGTACCTTGG CTCCCTTAG CCATGGCTGG GATGCAGGGC ACCAAGTCCT
28751 GAGACTGCAC AAAGCAGCAA GGCCCTGGGC CTGGCCCAGG AAACCATTTT
28801 TTCCCTCCTGG GCCTCTGGG CTATGATGGG AGGGCCCTTC CTGAAGACCT
28851 CTGAAGTGCC CTGGAGGCAT TTCCCCATT GTCTTAGTGA TTAACATTTC
28901 ACTCCTGTGTT TCTTATGCAG ATTCCTGCAG CTGGCTTGAA TTTTTCTC
28951 AGAAAATAGA TTTTCTTTT CTGTACATC ATCAGGGTGC AAATTGACA
29001 AACTTTGTGTC CTCTGCTTCC TGTTGAATGC TTTGCCACTT AGAAAATTCT
29051 TCTGCCTGAT ACCCCAAATC ATCTCTCTA GGTTCAAAGT TCCACAGATC
29101 TCTAGGGCAG GGGCAAAAGG CCACCAAGTCT CTTTGTCTATA GCATAACAAG
29151 AGTCATCTT GCTCCAGTTC CCAACAAGTT CCTCATCTCC ATCTGAGATC
29201 ATCTCAGCCT GGACTTCATT GCCCATATTA CTGTCAGCAT TTTGGTCAAA
29251 GCAATTCAAC AAGTCTCTGG GAACTTACAA ACTTTCCAC CTCTTTTGT
29301 CTTCTGAGCT CTCAAATT TTAAAGAAGTT CCAAACATTTC CCAGTCTTCT
29351 TCTGAACCTT CCTAACTGTGTT CCAACCTCTG CCTGTTACCC AGTTCAAAG
29401 TCAGTCCAT ATTGTTGGGT ATCCTTATAG TAGCACCCAA CTCTTAGTAC
29451 CAATTACTG TATTAGTTCA TTCTCACGCT GCTATAAAGA ACCACCTGAG
29501 AATGGGTATT TTATAAAGGA AAGAGGTTA ATTGACTCAC AGTTTCGCGT
29551 GGCTGGGGAG GCCTCAGATA ACTTACAGCC ATAGCAGAAA GGGAAAGCAA
29601 CATGTCCCTC ACATGGTGGC AGGAAGAAGA AGTGTGAGC AAAGAGGGAA
29651 AAGCCCTATA AAACCATCAT ATCTCGTGA AACTCACTCA CTATCATGAG
29701 AACAGCAGCA TGGGGTTGAC CACCCCCCAT AATTCAATT CCTCCACCA
29751 GCTGTCTCCC GTGACACATG GAAATTATGG GAACTACAAC TCAAGATGAG
29801 ATTGGGGTGG GGACACAGCC AAACCATATC ATCTAGGCTG GTATCGAAAT
29851 CCTGGGCTCA AGCAATCCAC CCACCTGCC CTACCAAAGT GCTGGGATTA
29901 CAGGCATGAG CCACCATATC TGAACGTGCT TTTGATTCT TTTGATTTTA
29951 ACCATCCATT GTTCTGCTT CTCTAGATAA CCCTGACTAA TATATAATTG
30001 GTATGAAGTG ATATCTCATG GCTTGATTT ATATTCTTT CATGGCTAGT
30051 GACTTTTTT GTACTTTGG GATATTGTTA TTATTATTAT TATTATTACT
30101 AGTGTTTATA CTTCTTCAGT AAAAGTGTAA GAAACAATT TTAAAGGCAG
30151 AATGTGACCA GAGTTCTGT TAGTTATATA ACCATCATGG ACCTTCCCTC
30201 AAGTGTAAAG CCATTAGTGT TACTCATGTC ACTCCAAATG TCAGCTTGT
30251 TTCTTCCTT TCACTGTCTC TTTGTGTCCC AAACCTGAAT TCATGGAAA
30301 AACATCTGAA TGGTGTTAA TATGGTTGG ATATTGTCC CCTCCAAATC
30351 TCATGTGAA ATATGACCTC CAGTGTGGG AGTAGGGACT ACTTGGGTCA
30401 CGAGAGTGGG TCCCTCATTA ATGGCTTGGT AATAAGTGA CTCTATTAGT
30451 TCATGAAAGC TGGTTGTTGA TAAGAGCTG GCATCTCATT TCTCTTGTCC
30501 TTCTCTCACC ATCTGACACA CTTGCTCACC TTTTTCTTC AGCCATGAGT
30551 AAAAGCTTCC TGAGGTCTCA CCAGAAACTG AGCAGATGTT GGTGCCATGC
30601 TTGTACAGTC TGAGAACTG TGAGCCAAT AAGCCTCTT TCTTTATAAA
30651 TTACCGAGTC TCAGGTGTTG GTTTAAAACA ACACAAAACA GACTAACACA
30701 GTGTTGATTG AAACAGCTGT GACTGGGTCA TCAGGGTGTAA AGAGAGGAGT
30751 CACTGAGTTG AAATATAGCC TCCTACTTAC ACCTGTTCAAG TAGAAGCTGT
30801 AGATATGAAG TAGCTGAAGC AGGCATTCCC TCTGAAACAT GTGTTTCACA
30851 TATGTCATAA TTATCTTCTG CTCTCATTT TCTTTAGGC TTTTGTCTCC
30901 ATCTCATTTC CCTGTTTAC TCTCATTTC ATATCTTAC ATTTCTTCT
30951 CCAGAATTGT TCAGAAGCTT GGAACCCCTC ACTCCAGTTA TTCTTTGACT
31001 ATGCAATTG TTTCTGTGCT TCATGGCACT TATGTTTGT ATCCTTGAC
31051 TTGTTGTTAT AGCTCAGTGG TTAGGAGTAC AGTTGGAGT TAGAATGCCT
31101 GGGTTGAAAC TCTTAATTCT ACTCTACTTA CTAGTCTTGT GACTATAACA
31151 AAATTCTTAG CCTCTCTTTG TCTGTAAAT GGAGAGTATA GTAAATACAT
31201 GGGCTTGTGTT TAAGGATTAA ATGAGTTAAC ATGTGAAATA CTTAGAACAA
31251 TGCCTGGCAA ATGCTCAATG AATATTGAGT ATTGCTTGCT TTTGTTTAGT
31301 GCCATGCCCTG TTGTTCCAC TGAGGGCACA GACCAGTGT ATCTGGTTAA
31351 CAGTTCTATG TCCACCACGT TGCAATAATG GACTCTCAGA AAATATTGAA
31401 GAATATGTTA AAGAATGAGT AGAATTATGC TACTGAAAAG GGTGAGTGG
31451 AGGTAGGTAG GGGAAAGGAC ATATACAGCC CTGGAGGCAG CATATATGGG

FIGURE 3, page 10 of 42

31501 GAATGGGTCA CACAGTGTTC CTTGGTACTC TCTAGACCAC AGTGGGCCAC
31551 CTCTTAGCTA GTGGCCTATG GATTATTCGA GCAGTCTGTT GGAAACATCC
31601 ATGAATATGA TAATAATGAC CCATTTGTGG GTTCTAAGAA AAAGGACAAC
31651 TACAATACTA GACAATAATA GTATGTAAGT TAGGAGGGAA GGGGATGATT
31701 TGTATTAAC TGTCTAAAA TTCTTACCTT ATTTAGGATG ATGGGGTCAG
31751 ACATTAACCT TAGACTTTGT TATATATATG TGGTAAAATT TCAAGGTAAA
31801 CCATTGAAAC TGTAGTAGTT GAGTATATAA CTTCCAAATC AGGGGGAAA
31851 GAAATGGAAT AAGAAAATAA ATACATAAAC ATAAGATTGA AACAACTCAA
31901 TGAAGAGTAG AGAGAAGAGG GAAAACATA GAAAGAATGA GATAATTAGA
31951 AAGCAATAGG TAAGATGTGA GAAAATAATT CAAGTACAGT AAAACTCCAC
32001 TAAAATGTGC CCTGCAGTAA TGTTGGGCA TGATTCCCT TCATCCCCAT
32051 TCTCAAATGG GGCAGCCTAA ATAGCGTTCT TATCCTGTTT CCCTGGGGGT
32101 TTGAGGTGGG TGACGAGTAA GTTAGAAGAT AATCACCTTC TGATCAGTTA
32151 GGACTTTCTC AGTTAGTCT TCAATTAATA AAAATAATG TAAATTCAT
32201 CAGAAGGCAG AGATTGTCAG ATGAAAGAAC AAGCAAAATA AAAGTCTTAC
32251 TGAAAAAAAG CTGGGGTAGC TATGTTAATA TCAACTGTT ATTATTATTA
32301 ATAATCTATT AATAATAGAT TATATAGTAA AAACATTAAT AAAAATAGAG
32351 TGTCACTACA TTTAAAATT CAGTATGAGG ATATACAATT TTTAAGCTGG
32401 TTGATAAAAT TCTGGGATT AATTGGCAAA TCCATCATAG TGGTGAGAGA
32451 TTTTAACACA ATTCTTCCTG TATTTGATAG GTCAAGCAGA GAAAAACTTT
32501 AGTGAAGACA AAAACTTCTA AATACATAAG CTTGATTTAA TGGGCATGTA
32551 ATAGGACCTA GCATCAAAAA ATTAGAAAAA ATATTTTTTCTC TTAGGTATTT
32601 ATGGAACATG TATAAAAATT GATTCGTAG TAGGCCATAA AGCCAGGTT
32651 AACACATTTCA AAAGAACTGG TATCACAAGA ACTGCTTCT CTGACCACTA
32701 TGCATTAATAA TAGAAGTTAA TTACAGACAT AAATTATAAA AATGCCAATA
32751 TTTTAAAGTG TGATATACAC TTCTCAACTT ATGGTCAAA GGAAATCGTA
32801 AGTGGAAATT CAAGGACACG TTGACTTGAA AACATTAATAA CTTATGGAAT
32851 ATTTCTAAGA TGGAACCTGT ATGAATTAA TAGTCTGAAA GCTTTTATTA
32901 GAAAAGAATT AAGTCTGAAA ATTAATGTGC TAAGTTAGGG GAGAGAAAAT
32951 GGAATAATCT CGAAGAAGGT AGGAGGAAG AGATAATAAA GAATATATAG
33001 CAAAGATGCA GTAACAGGAT CAACAAAGCC AGAAACTGTT GGAAAAGACA
33051 AGCCTCTGGA AAGATTGATG AAGAAAAAAAG AGAAATGAGA TGTAATAAA
33101 TCATGTTCAAG TTATAAATAG GCACATAAGG ACTTTAAAAA AACTAATAAA
33151 ATAATATGAA TCATTAATGC CAATAAATTG GAAAACAGAC AAAGTAGGTG
33201 AATTCTAGA AAAATATAAC TTACTGGGAC TGAATGAAGA AGCAACAGCT
33251 TATAGTACCT AAGCAATTGA AGAGATTGGG TCAGTAATTAA AAAATTTCT
33301 CATAAACAAA ACGTTAGCCC CAGATGGTTC TTGCAAATGA TAAAGAAC
33351 GATGTACAAA CATTCCAGA GTGTAGAAGT ACACTGTCCT ATCCTTCTA
33401 GGAGATCATT ATAACACCAA AAGCAGACAG TATATGAAAC AGGGAAATT
33451 GAGGCCAAGA TACCTATGAC TTATATGTAA AAATTAAAG AAAATATTAG
33501 CAAACTGAAT CAGCCATTAA AAAAAATATA CCACAATCAA TGCATTCTA
33551 AGAGCAGCTT AACAAAATTG GTTAGAAGGC ATTAAGAAG ACTCAGTATA
33601 GAAAAGATGT ACCTTCTCTC CAAATTGGT ATAGAGATTC AATGCCATTA
33651 AAAAACCCCA CCTGGTTTTT TTGAGGAAC TGTCAAGCTG AGTCTCAAAT
33701 TTATATCAAA GAGCAAAGGC CTAAGAATAT CCAGGACATT CCTGAAGAAC
33751 TGTAAGGAGC CAGGGGCTG CCCTATCAGA TACCAAGGGT TGTTATTAAG
33801 CCATAACCAA GTCAGTGCTG TTTCTACAGA AACAGACAAG TTAACAAGTG
33851 AAACATAATA GAGAGCCCAG AAACAGACCC ATCCATATTG TGGATTTGTC
33901 ACGTGAAAGA AGTAGCTTTG CAAAACTTG GGAAAAGGAG AGTGTGTGCA
33951 ATAGATGATG CTCGTGCTCA TGCAAGACAA AAGGAAATTG GGATACCTGC
34001 CTCTTACCGT ACACAAACAC CAACCTAAAC GTGAAAGTT AACTATAACA
34051 GCTTGAGGTG GTGGGGAAAGA AATATCTTA TCTCAGTGTAA GGGAAAGATT
34101 TATTTAAAAA AGAAGACACA AAAGGCCATA CATAGGAATG AAAAGATTGA
34151 ATTCAAGCTGC ATTAATTAAGA TTAAATTCTAG CTGCGTTAAATCAGAAG
34201 TCTGTACTTG GACAGCATAG AGTGGAAAGA CAAAGAGAAG GTATTTGCCA
34251 GCTTATAACT TGAAGGATTA GAATGAATGA TATAAGAAC TATGTAATAA
34301 AGAAAAAGAC ATACAACCGG TTAGAAAAAC GGGCAAAGAC ATGAACAGCA
34351 TATTTCACGT GAAGGAAACA GCGGTAGCAA ATGAACATGG TAAGAGATGC
34401 TCAACACGTT TAGTAATTG AAGGGAAATG CAAGTTATAAC CCACAGCAAG
34451 ACTATCTTAT CTAGGAAGTT TGTCAATACC CTAATGTTC TGTGGTTTA
34501 AGCTACAGAG TTGTAATTCA ATTTATTAT TCAATAATAA CTCAGTGGCA
34551 GGCACTGTT TAGAACCTT GGTTATAACT TTGAATGAAA TTAAAAAAA
34601 TCCTTGCCTT GTGGAGGATG CTTATGTGTG GGGAGTTGGG TGGTGGGGTC

FIGURE 3, page 11 of 42

34651 AAACAACAAT TACATTAAAA TAGAAAATAG TGACATAAAT AACCTATAA
34701 ATATTGCAAC CCAGAGTTAT ATTATAAATG TAAGTAGTGA CTAGGACTCT
34751 CATGCAGATA TACCTCTGTG CTGGGACAAA TGAAAGTTA AGTGTAAATT
34801 CCCATATGCA AGTCAAAATA AAAAGTGACA CTAGAAAACA CAATAATGAA
34851 TATCTGAAAA TTGCATTTA TTTGACTGCC ATCCTTTGC ATCATTTCA
34901 TACTAATTAT AGAATAAAAT TTGTAAGGATG CACCAAAGCT TTTTTAGAG
34951 ACATCCATTA ATTCAATAAA TAAATGAGCA CCTTCTTGT GCCAGCAGCT
35001 GTAAGAGGTG GCCCAAGGAA GGGAAATAAAA CAGTCAAAAT CCTGGTACAC
35051 TCAGAGTTTC TCTTAGGAGA AAACAGATAC AAATGGCATT AATTACCAAG
35101 AAACTTGTAACAAAGCCAA ATATAATGA TAAATATTG AGTACAGTAT
35151 GTTAATTTTA AGATTGAAAA TGAGGTGCCA GGATTCTTA AGACTCAAAG
35201 GCGAAGATGG CTGAATAGGA ACAGCTCTGG TCTACAGCTC CCAGCGTGAG
35251 CGACGCAGAA GACGCATGAT TGCTGCATT CCATCTGAGG TACCGGGTTC
35301 ATCTCACTAG GGAGTGCCAG ACAGTGGGCG CAGTCAGTG GGTGTGTGCA
35351 CCGTGCAGA GCTGAAGCAG GGCAGGGCAT TGCCTCACTC GGGAAAGTCCA
35401 AGGGGTCAGG GAGTCCCTT TCCTAGTCAA AGAAAGGGGT GACAGATGGC
35451 ACCTGGAAAA TCGGGTCACT CCCACCTGAA TACTGCACTT TTCTGACGGG
35501 CTTAAAAAAT GGCGCACCA GAGATTATAT CCTGCACCTG GCTCGGAGGG
35551 TCCTACACCC ACGGAGTCTC GCTGATTGCT AGCACAGCAG TCTGAGATCA
35601 AACTGCAAGG CGCGCGCAG GCTGGGGGAG GGGCACCCGC CATTGCCAG
35651 GCTTGCTTAG GTAAACAAAG CAGCCGGGAA GCTCAAACGT GGTGGAGCCC
35701 ACCACAGCTC AAGGAGGCCT GCCTGCCTCT GTAGGCTCCA CCTCTGGGGG
35751 CAGGGCACAG ACAAACAAAAA AGACAGCAGT AACCTCTGCA GACTTAAATG
35801 TCCCTGCTG ACAGCTTGA AGAGAGCAGT GGTTCTCCC GCACGCAGCT
35851 GGAGATCTGA GAACGGGCAG ACTGCCTCCT CAAGTGGGTC CCTGACCCCT
35901 GACGCCGAG CAGCCTAACT GGGAGGCACC CCCCAGCAGG GGCACACTGA
35951 CACCTCACAC AGCCGGTTAC TCCAACAGAC CTGCAGCTGA GGGTCCTGTC
36001 TGTTAGAAGG AAAACTAACAA AACAGAAAAG ACATCCACAC CAAAAACCCA
36051 TCTGTACATC ACCATCATCA AAGACCAAAA GTAGATAAAA CCACAAAGAT
36101 GGGGAAAAAA CAGAGCAGAA AAACGGAAA CTCTAAAAAG CAGAGTGCCT
36151 CTCCTCCTCC AAAGGAACGC TGTTCCCTAC CAGCAACGGA ACAAAAGCTGG
36201 ATGGAGAATG ACTCTGACGA GCTGAGAGAA GGCTTCAGAC GATCAAATT
36251 CTCTGAGCTA TGGGAGGACA TTCAAACCAA AGGCAAAGAA TTGAAAACT
36301 TTGAAAAAAA TGTAGAAGAA TGTATAACTA GAATAACCAA TACAGAGAA
36351 TGCTTAAAGG AGCTGATGGA GCTGAAAACC AAGGCTCGAG AACTACATGA
36401 AGAATCCAGA AGCCTCAGGA GCTGATGCGA TCAACTGGAA GAAAGGGTAT
36451 CAGCGATGGA AGATGAAATG AATGAAATGA AGCGAGAAGG GAAGTTTAGA
36501 GAAAAAAAGAA TAAAAAGAAA CGAGCAAAGC CTCCAAGAAA TATGGGACTA
36551 TGTGAAAAGA CCAAATCTAT GTCTGATTGG TGTACCTGAA AGTGACGGGG
36601 AGAATGGAAC CAAGTTGGAA AACACTCTGC AGGATATTAT CCAGGAGAAC
36651 TTCCCCAATC TAGCAAGGCA GGCCAACATT CAGATTCAAGG AAATACAGAG
36701 AACGCCACAA AGATACTCCT TGAGAAGAGC AACTCCAAGA CACATAATTG
36751 TCAGATTCAAC CAAAGTTGAA ATGAAGGGAA AAATGTTAAG GGCAGCCAGA
36801 GAGAAAAGTC GGGTTACCCCT CAAATGGAAG CCCATCAGAC TAACAGCGGA
36851 TCTCTGGCA GAAACTCTAC AAACCAGAAAG AGAGTGGGGG CCAATATTCA
36901 ACATTCTAA AGAAAAGAAT TTTCACCCCA GAATTCATA TCCAGCCAAA
36951 CTAAGCTTC TAAGTGAAGG AGAAATAAAA TCCTTTACAG ACAAGCAAAT
37001 GCTGAGAGAT TTTGTCACCA CCAGGCCTGC CCTAAAAGAG TTCTGAAAGG
37051 AAGTGTCTAA CTTGGAAAGG AACAATCAGT ACCAGCCGCT GCAAATCAT
37101 GCCAAAATGT AAAGACCGTC GAGACTAGGA AGAAAATGCA TTAACAAACG
37151 AGCAAAATAA CCAGCTAACCA TCATAATGAC AGGATCAAAT TCACACATAA
37201 CAATATTAAC TTAAATGTA AATGGACTAA ATGCTCCAAT TGAAAGACAC
37251 AGACTGGCAA ATTGGATACA GAGTCAAGAC CCATCAGTGT GCTGTATTAA
37301 GGAAACCCAT CTCACATGTA GAGACACACA TAGGCTAAA ATAAAAGGAT
37351 GGAGGAAGAT CTACCAAGCA AATGGAAAAC AAAAAGAC AGGGGTTGCA
37401 ATCCTAGTCT CTGATAAAAC AGACTTTAA CCAACAAAGA TCAGAAGAGA
37451 CAAAGAAGGC CATTACATAA TGGTAAAGGG ATCAATTCAA CAAGAAGAGC
37501 TAACTATCCT AAATATATAT GCACCCATAA CAGGAGCACC CAGATTCAA
37551 AAGCAAGTCC TGAGTGACCT ACAAAAGAGAC TTAAACTCCC ACACATTAAT
37601 AATGGGAGAC TTTCACACCC CACTGTCAAC ATTAGACAGA CCAATGAGAC
37651 AGAAAAGTCAA CAAGGATACC CAGGAATTGA ACTCAGCTCT GCACCAAGCA
37701 GACCTAATAC ACATCTACAG AACTCTGCAC CCCAATCAA CAGAATATAC
37751 ATTTTTTCA GCACCCACACC ACGGCTATTC CAAAATTGAC CACATACTTG

FIGURE 3, page 12 of 42

37801 GAAGTAAAGC ACTCCTCAC C AAATGTAAAAA GAACAGAAAAT TATAGCAAAC
37851 TATCTCTAG ACCACAGTGC AATCAAAC TA GAACTCAGGA TTAAGAATCT
37901 CACTCAAAC CGCTCAACTA CATGGAAACT GAACAAACCTG CTCCCTGAATG
37951 ACTACTGGGT ACATAACGAA ATGAAGGCAG AAATAAAGAC GCTCTTGAA
38001 ACCAACAAAGA ACAAAAGACAC AACATACCG AATCTCTGGG ACGCATTCAA
38051 AGCAGTGTGT AGAGGGAAAT TTATAGCACT AAATGCCAC AAGAGAAAAGC
38101 AGGAAAGATC CAAATTGAC ACCCTAACAT CACAATTAAA AGAACTAGAA
38151 AAGCAAGAGC AAACACATT AAAAGCTAGC AGAAGGCAAG AAATAACTAA
38201 AATCAGAGCA GAACTGAAGG AAATAGAGAC ACAAAAAACC CTTCAAAAAA
38251 TTAATGAATC CAGGAGCTGG TTGTTTTGA AAGGATCAAC AAAATTGATA
38301 GACCGCTAGC AAGACTAATA AAGAAAAAAA GAGAGAAGAA TCAAATAGAC
38351 ACAATAAAA ATGATAAAGG GGATATCACC ACCAATCCC CAGAAATACA
38401 AACTACCATC AGAGAATAC ACAAACACCT CTATGCAAAT AACTAGAAA
38451 ATCTAGAAGA AATGGATAAA TTCTCGACA CATACACCC CCCCAGACTA
38501 AACCAAGAG AAGTTGAATT TCTGAATAGA CCAATAACAG GATCTGAAAT
38551 TGTGGCAATA ATCAATAGCT TACCAACCAA AAAGAGTCCA GGACCAGATG
38601 GATTCAAGC CGAATTCTAC CAGAGGTACA AGGAGGAAC GGTACCATTC
38651 CTTCTGAAAC TATTCATC AATAGAAAAA GAGGGATCC TCCCTAATC
38701 ATTATATGAG GCCAGCATCA TCCTGATACC AAAGCCAGGC AGAGACACAA
38751 CAAAAAAAGA GAATTTTAGA CCAATATCCT TGATGAACAT TGATGCAAA
38801 ATCCTCAATA AAATACTGGC AAACTGAATC CAGCAGCACA TCAAAAAGCT
38851 TATCCACCAT GATCAAGTGG GCTTCATCCC TGGGATGCAA GGCTGGTTCA
38901 ATATAACGCAA ATCAAGTAAAT GTAATCCAGC ATATAAACAG ACCAAAGAC
38951 AAAAACCCACA TGATTATCTC AATAGATGCA GAAAAAGCCT TTGACAAAAT
39001 TCAACAAACAC TTCATGCTAA AAACCTTCAA TAAATTAGGT ATTGATGGGA
39051 TGTATCTCAA AATAATAACA GCTATCTATG ACAAAACCCAC AGCCAATATC
39101 ATACTGACTG GGTAAAAACT GGAAGCATT C CTTGAAAAA CTGGCACAAG
39151 ACAGGGATGC CCTCTCTCAC CACTCCTATT CGACATAGTG TTGGAAGTTC
39201 TGGCCAGGGC AGTTAGGCAG GAGAAGGAAA TAAAGGGTAT TCAATTAGGA
39251 AAAGAGGAAG TCAAATTGTC CCTGTTGCA GACGACATGA TTGTATATCT
39301 AGAAAACCCC ATTGTCTCAG CCCAAAATCT CCTTAAGCTG ATAAGCAACT
39351 TCAGCAAAGT CTCAGGATAC AAAATCAATG TACAAAATC ACAAGCATT
39401 TTATACACCA GCAACAGACA GAGAGCCAA TCATGAGTGA ACTCCCGTTC
39451 ACAATTGCTA CAAAGAGAAAT AAAATACCTA GGAATCCAAC TTACAAGGG
39501 TGTGAAGGAC CTCTTCAAGG AGAAGTGC CAA ACCACTGCTT AATGAAAATA
39551 AAGAGGATAC AAACAAATGG AAGAACATTC CATGCTCATG GGTAGGAAGA
39601 ATCAGTATCG TGAAAATGGC CATACTGCC AAGGCAATT ACAGATTCAA
39651 TGCCATCCCC ATCAAGCTAC CAATGACTT CTTCACAGAA TTGGAAAAAA
39701 CTACTTTAA GTTCAATATGG AACCAAAAAA GAGCCCGCAT TGCCAAGTCA
39751 ATCCTAAAGCC AAAAGAACAA AGCTGGAGGC ATCATGCTAC CTGACTTCAA
39801 ACTATACACT AAGGCTACAG TAACCAAACC AGCATGGTAC TGGTACCAAA
39851 ACAGAGATAT AGACCAATGG AACAGAACAG AGCCCTCAGA ATAACGCCG
39901 CACATCTACA ACTATCTGAT CTTTGACAAA CCTGAGAAAA ACAAGCAATG
39951 GGGAAAGGAT TCCCTATTAA ATAAATGGTG CTGGAAAAC TGGCTAGCCA
40001 TATGTAGAAA GCTGAAACTG GATCCCTTCC TTACACCTTA TACAAAATC
40051 AATTCAAGAT GGATTAAAGA CTTAAACGTT AGACCTAAA CCATAAAACC
40101 CCTAGAAGAA AACCTAGGCA TTACCATCA GGACATAGGC ATGGGCAAGG
40151 ACTTCATGTC TAAAACACCA AAAGCAATGG CAACAAAAGC CAAAATTGAC
40201 AAATGGGATC TAATTAAACT AAAGAGCTTC TGCACAGCAA AAGAAACTAC
40251 TATCAGAGTG AACAGGCAAC CTCCAAAATG GGAGAAAATT TTTGCAACCT
40301 ACTCATCTGA CAAAGGGCTA ATATCCAGAA TCTACAATGA ACTCAAACAA
40351 ATTTACAAGA AAAAAGACAA ACAACCCAT CAAAAAGTGG GTGAAGGACA
40401 TGAACAGACA CTTCTCGAAA GAAGACATTG ATGCAGCCAA AAAACACATG
40451 AAAAATGCT CACCATCACT GGCCATCAGA GAAATGCAA TCAAAACCC
40501 AATGAGATAC CATCTCACAC CAGTTAGAAT GGCAATCATT AAAAAGTCAG
40551 GAAACAAACAG GTGCTGGAGA GGATGTGGAG AAATAGGAAC ACTTTTACAC
40601 TGTTGGTGGG ACTGTAAACT AGTTCAACCC TTGTTGGAAGT CAGTGTGGCA
40651 ATTCCTCAGG GATCTAGAAC TAGAAATATC ATTTGACCCCA GCCATCCCAT
40701 TACTGGGTAT ATACCCAAAG GACTATAAT CATGCTGCTA TAAAGACACA
40751 TGCACATGTA TGTTTATTGT GGCACTATTG ACAATAGCAA AGACTTGGAA
40801 CCAAGCCAAA TGTCACCAA TGATAGACTG GATTAAGAAA ATGTGGCACA
40851 TTTACACCAT GGAATACTAT GCAGCCATAA AAGATGAGTT CATGTCTTT
40901 GTAGGGACAT GGATGAAATT GGAAATCATC ATTCTCAGTA AACTATCACA

FIGURE 3, page 13 of 42

10
9
8
7
6
5
4
3
2
1

40951 AGAACAAAAA ACCAACACCC GCATATTCTC ACTCATAGGT GGGAAATTGAA
41001 CAGTGAGAAC ACATGGACAC AGGAAGGGGA ACATCACACT CTGGGGACTG
41051 TTGTGGGTG GGGGGAGGGG GAGGGATGGC ATTGGGAGAT ATACCTAATG
41101 CTAGATGACG AGTTAGTGGG TGCAGCGCAC CAGCAAGGCA CATGTATA
41151 TATGTAACTA ACCTGCACAT TGTGCACATG TACCTAAAAA CTTAAAGTAT
41201 AATAATAAAA AAAAAGACT CAAAGGCACA GTCACTGACA GTTTGATTT
41251 TTATAATAGC TGTTAATTCTC CCTAACTTCG AGGAAGTTGA TAGCATGTT
41301 TGAGTATATT TCAAAACTAC ATTCAAATGT TGCAATAGAA CATTAAGAAT
41351 TATCTTCATG ATCCACTAAG TGCATGAAAA AAATGGATAA TGAATCTATT
41401 CATTACCATC GTTTAATATT TTATCTTCAA GTTTTGTGT TTTGTAGCTC
41451 ATTGGCAGAG TTTGACAGAG TGCTGAAAGT ATTCTTAGT GAGCTGGCTG
41501 TAATTTTGG GCCCATTTTT ATCTAGATAA TTAAAACATAT CTGACAGGAC
41551 CATAAAATGC TTGCTGCCAT TTCCAACAC CTATATTGT GGATGGGTT
41601 TTTTAATTAA ATGAGAATAT TATGTTAGAA AAGAAACTGT CATTCTGTAA
41651 AGTGGCAAT AATGTTAGTT TTATTTATCA ATTTAGTTT GTACTTTGAT
41701 CATTTTTTA AAATTCAGC ATTGATGTTG ATGGGACAAT GACAGTGGAC
41751 TGGAAATGAAT GGAGAGACTA CTTCTTATTT AATCTGTAA CAGACATTGA
41801 GGAAATATC CGTTCTGGA AACATTCTAC AGTAAGTCTA CTTTATGTAT
41851 TTATACTTAT TTGGAGCTAT AAACCATAGG TACAGTTATC ACCCAAGAAC
41901 ACTCTGTAAC ACTTATGGC CAGGATAACCT GAGTCCCAGT AGCTCCTTAA
41951 CCTGTAGAGT TCTATTTATT CTATTAGGCA TAGATTATA GAGTATTAAA
42001 CAAAAAAA CAGCTCTCCC TCTCCCTCTC CCTCTCTCTC CCCCTCCCCA
42051 CGGTCTCCCT CFCCCTCTCT TTCCACGGTC TCCCTCTGAT GCGGAGCCAA
42101 AGCTGGACTG TACTGCTGCC ATCTCGGCTC ACTGCAACCT CCCTGCCTGA
42151 TTCTCCGTGC TCAGCCTGCC GAGTGCCTGC GATTGCAGGC GCGCACCGCC
42201 ACGCCTGACT GTTTTCGTA TTTTTTGGT GGAGACGGGG TTTCGCTATG
42251 TTGGCCGGGC TGGTCTCCAG CTCCCTGACCG CGAGTGATCC ACCAGCCTCG
42301 GCCTCCCGAG GTGCTGGGAT TGCAGACGGA GTCTCGTTCA CTCAGTGCTC
42351 AATGGTGCCTC AGGCTGGGGT GCAGTGGCAT GATCTCGGCT CGCTACAACC
42401 TCCACCTCCC AGCCGCCTGC CTTGGCCTCC CAAAGTGCCTA AGATTGCAGC
42451 CTCTGCCAG CGGCCACCCC GTCTGGGAAAG TGAGGAGCGT CTCTGCCTGG
42501 CCGCCCATCG TCTGGGATAT GAGGAGCCCC TCTGCCTGGC TGCCCAGTCT
42551 GGAAAGTGGAG GAGTGTCTCT GCCCCGGCCGC CATCCTGTCT AGGAAGTGGAG
42601 CGTCTGTCGC CGGCCGCCCA TCGTCTGGGAG TGTGAGGAGC CCCTCTGCCT
42651 GGCTGCCAG TCTGGAAAGT GAGGAGCGCC TCTTCCCGGC CGCCATCCCC
42701 TCTAGGAAGT GAGGAGCGTC TCTGCCCGC CGCCCATCGT CTGAGATGTG
42751 GGGAGCCCT CTGCCCGGCC GCCCCGTCTG GGATGTGAGG AGCGCCTCTG
42801 CTCGGCCGCC CCGTCTGAGA AGTGAGGAGA CCCTCCGCC GGCAGCCGCC
42851 CCGTCTGGGA AGTGGAGGAGC GTCTCCGCC GGCAGCCACCC CTGTCCGGGA
42901 GGGAGGTGGA GGGGTCAAGCC CCCCAGCCCC CCAGCCACCC CATCCGGGAG
42951 GTGAGGGGTG CCTCTGCCCG GCGCCCTCTA CAGGGAAGTG AGGAGCCCC
43001 CTGCCCGGCC ACCACCCCAT CTGGGAGGTG TACCCAACAG CTCATTGAGA
43051 ACGGGCCATG ATGACAATGG CGGTTTGTG GAATAGAAAA AGGGGAGAGG
43101 TGGGAAAAG ATTGAGAAAT CGGATGGTTG CTGTGTCTGT GTAGAAAGAG
43151 GTAGACATGG GAGACTTTTC ATTTTGTCT GTACTAAGAA AAATTCTTCT
43201 GCCTGGGAT CCTGTTGATC TATGACCTTA CCCCCAACCC TGTGCTCTCT
43251 GAAACATGTG CTGTGTCCAC TCAGGGTTAA ATGGATTAAG GGCAGGTGCAA
43301 GATGTGCTTT GCTAACAGA TGCTTGAAGG CAGCAGGCTC GTTAAGAGTC
43351 ATCACCAACTC CCTAATCTCA AGTACCCAGG GACACAAACCA CTGCGGAAGG
43401 CCGCAGGGTC CTCTGCCTAG GAAAACCAGA GACCTTGTGTT CACTTGTAA
43451 TCTGCTGACC TCCCCTCCAC TATTGTCTG TGACCCCTGCC AAATCCCCCT
43501 CTGCGAGAAA CACCCAAAGAA TGATCAATTA AAAAAAAA AAAAAAAACA
43551 ACCCAAGACT GCATAAAATGT CCATTCTGAA AACTTGGAAAG AAGTACCA
43601 TTGATGAATA AGCTGTCTAG CTTTTATTGG CATTAAAGTA TTCTGCCATA
43651 GGGAAAGTGTAA AAAGTTGTAG GCTTTACTT TTTATAGGTAA CTATATTGTC
43701 CAAATAATCT CAGCACCTCA TGGTTGCTAA GGATCTGTGT CTTGTTGG
43751 TCAGATTATG TTTATCTCTG GCATAAGGCA CTTAACAAATA TTCATTAAAG
43801 GTTACAGAAT CTTTTGCTT CATCTGCTTA GCATTTCATA CCAGTTGT
43851 TTCCACCAAA CTTCAAAATT TTGATTGTGTT CATTAAATTCTGCATAACTG
43901 ATGTAAACCA AGTTCTATTAA TTGTGCAATC TGCTCCTGAA ACCCTTAGGA
43951 ACTCTCTGAA GGAGTTTTAT TTATTTTTG TTTTTGTTTT TGTTTTGTT
44001 TTGTTTTTT GAGACGGAGT CTTGCTCTGT TGCCCAAGGCT AGAGTGCAGT
44051 GGTGCGATCT CGGCTCTCTG CAAACTCGGC CTCCGGGGTT CACGCCATT

44101 TCCTGCCTCA GCCACCGGAG TAGCTGGAC TACAGGCACC CACCACTGCG
44151 CCTGGCTAAT TTTTTTGTA TTTTTAGTAG AGACGGGTT TCACCGTGT
44201 AGCCAGGATG GTCTCGATCT CCTGACCTG TAATCCGCC GCCTCGCCTC
44251 CCAAAGTGCT GGGATTACAG GCGTGAGCCA CTGTGCCCGG CCTTTTTTT
44301 TTTTTTTCT TTATGGGCTT GTCTTCTACA CTTCAGATTT GACTAAATTA
44351 AATATGCATT AAATGAAGTC AGGAGTTCAC ATTGCCACTA GTAACAATGC
44401 CTAAGCTTAC ATAAGCATT ATAAAATTGT TGGTATTAG TGCCCTCTCA
44451 GCTATGAGTA TAAGATAATA TTATACAGT AGTTCACTTG CCTAGATAAA
44501 TTGTACACTA TGTGAAGTT TATTTACATA ATTCTACGG TATTTTTAA
44551 GGTAGTTGAT AACAGTTGAG ACTACAATTG TATCTCATT TTATTGATAG
44601 TAAAATGAAG GAAGGGAGGG TTACTACCAT AGGAGAGCTC CTCCCCGTTG
44651 CACTCTGCC TGAAAAAATT TTTCTGCCAA AACAAATTAG ATAATAGAAT
44701 TGTAATAATA TTATTATAGA ATTGTTCTC TCAAACATA GTAATGTAGA
44751 ATAGGTGAA GGGGTGATGA TTGAAACAA TACCTCTCCA TTAGCTAAAT
44801 TTTATATAGA ATCTATTGCA TGTTTAAAT GATAAGTCAG ATTTATAAAA
44851 ATATTTTAT AACAGTAGG AAATGAGTT AGGGTATTG ACATACAGTT
44901 TTAATTTTA TTTACATATT TAAAACATAT CATGGTATAA ATATGATGTG
44951 GATATAAATT TGAGATAAAG GAAAGTATTG TTAAGAATTG ATGAACATAAT
45001 TTCTAAAAG ATGTCATCAC CAGTTGGTT TCTAGCCTTA TGAAAAATGG
45051 TTGCAATAAA AAAGATTGAC TATGATAAAA TGCTGCCCTT TCATTTAAC
45101 CTAGACCAAG AGAAAACATA CTGTGAATCT ATGATGAATG AAAGAAAGTT
45151 GTAACTGTTG GTTTGTTATA TTGTAATTA CTGTTTATTG TCATTTCTTG
45201 TGAACTGATA CTGTAACCTTG TTCATTGTGA TAGACAACT TATAATCTAT
45251 GTACTCAAAT TGGTTTAGTA TAAATTCTAG GGAATGAAGT TCATATTAAC
45301 TGTAATAATA CATGATTGTT CTCTAAAACA AAACGTCTTC TGGGATTATT
45351 TTTAACTAAG GCGCATGGGG ATCTTTTTT CATTTTACA GGAAATTGAC
45401 ATAGGGGATA GCTTAACATAT TCCAGATGAA TTCACCGAAG ACGAAAAAAA
45451 ATCCGGACAA TGGTGGAGGC AGCTTTGGC AGGAGGCATT GCTGGTGCTG
45501 TCTCTCGAAC AAGCACTGCC CCTTTGGACC GTCTGAAAAT CATGATGCAG
45551 GTGAGCTTAA TTATCGTGTG TCCAGGTTTG CCCTAAATAT TCTAAACAA
45601 TGAGAAATGT GGTGTTGAA AAAAGAAGTT TTAAATTTT TCAGTAATAA
45651 TCTTTTATAC CCTAAAAAAT AAATCTATT TGTTGCTGTT AACTCTAAAT
45701 TCAGTCCATG TAAGTATGGC AGTGTACCAA ACCTTAAATT GTTAGTACAT
45751 GTGTGTAATG AACTTTTAAT CTTTGGCATT CTATGACTAT TCAAACATT
45801 AATTCAAAAA ATATCTCTAG CTATTGTTG AGGATTCTCC TGATTTATAG
45851 TTTCCCTCTT TTAAATATAC TTATCATAAA GTAAAGTATT TTGAAATCT
45901 AGACTCTTAG AGCAGCAATG TAATTTGAA AATTATTCTA AAGCTGAGGT
45951 TAGCAGAAAA AGATCTGGCT TTATAGACTG ACTTTGCTAT TTACTAGCAG
46001 TGTAGCATTG GGCTGGCCAG AGTGGAAAGA GGGAAATGGAA AAGAATTAAAT
46051 ATGTATTGTC TCACTGTGGT AACCCAGTTA ATCCTTGCAG CAGCCCAGTG
46101 AAGTAGGTAT TTATCATT TTCCAGGGGG AATCTGAGGC CCAGAGAATT
46151 GACTTTTCCT TTACAACAAA TGAGAGGGGG AATGCAGTAT CTTTGCCTCC
46201 AGTGCCTCTG GTTCTCATGC TGCATGAAAC CTCTGAGGTC TCATTTTCT
46251 TCATTCGGG ATGGGATAA GAATATCTAA TAAGAATGGT TTAAGAATCA
46301 AGCAATATCA GGTATGTGAT AATGTCTGGT ACACGGAAT AACCTATTGG
46351 AACATAGTAG TTGTTTACAA AATATTTTA AAACCTTGTT ATACTTATGG
46401 TCAACACTTT TTATATTGCT CTGTAGATTT CTGTACAAAA AGATTCTGAC
46451 ACTGTTTAA GCCAGCATTG CTTCAGAATG TACCCAAATC TCAAAATTAA
46501 TTTAGGGCA AAGCTAATGC TTAAAGAAA AAGGAGAGGG GATTGGTG
46551 TGTTTTCTT TAGGAACAGT AGTAACCTGA CTTTAGAGA ACTTGAATAA
46601 GCATTATTT TTCTCTTGT CCTATTTAT TGTGAAGTTT ATTTATTTAA
46651 AATAAAATGG ATTCTCTGG AATTAGTTT CTGCAAATTG GAGGAGTTTC
46701 CAAAGTCAC CTTCAGGTT GATACTCTC TAGAAAGACT CACATAACTC
46751 ACTGAAAGCT TATTACCCCT GGTTATGGTT TATTACGGGG AAAAGATGCG
46801 GATGAAAATC AGTCAAGTAA AGAACACAT AGGGCAGAGC TTCTGTTGTC
46851 CTCTCCCTGT GGAGTCTCCA TGTCTTACTT TCCTGGCACT GTTATGTGGC
46901 ACTAGGCATG GAATATTGCA GACCAACCG GGAAGCTCAC CTGAGCCTT
46951 GGTGTGCAGA GTTCTTATTG GGGCCTGTT TCATACTGGC CACATGGCTG
47001 GCCTTCAGAA TTCAACCGT TCTGTGAGTG TGTGTGTGTG TGTGTGTGTG
47051 TGTGTGTGTG TTGTTAGTGG TAGTCACCC TTATGTGA GCTGAAACAA
47101 TCAGAAGAAT AGCTGATTTG TTAAATTATT TTGGGTGTAT TGGACTTAAT
47151 CAGTTTTAT CTGTAGGTGG TCATAAGGTA CAGTATTAA AAGTGAATAC
47201 CACATCTGTA GTATAAGCCA AGTAATTAT CAGTACTCAC AGGATGGTA

FIGURE 3, page 15 of 42

0 1 2 3 4 5 6 7 8 9 A T C G

47251 CATGTTGTAA TGAATTATT GCCTAGAGAG GGCCTCAAAA TATGCCAAG
47301 AGGGTGCAAT TTTTATTTT GGTTTCAGGC TGTATGCATT CCAGTGTGG
47351 TAGCCCTGAT ATACACAATA TCCAAACCAT TTCAGACCCA TTTACAGTTC
47401 ATGTCTGTAC TACTTCTTGA GGAGAGGGAG TAACATATTA CTTTAAATTA
47451 TATGTAATAA TATACATACA TTAAATTATA TGTAATAATA TAATATTATT
47501 ATTTGCAGTA TACTTTTTA TTTCCTTTA ACTGAGCTTG TTCACTGTTTC
47551 AAAGGGTGT CCATTGCCTG ATACATATT TAGTTAATAT TATCTTATGA
47601 AGGTTGTTCA TAATTTAAT ACTCTCTTG TCTTCTCTCT CTGCTTCTC
47651 ACACTGAAGA TACCAATTAT TCTTAGTTT AGAGTCAGAG ACAGGCCTCT
47701 AAAATCATGG CAATACTCCC TCTCATGATT ATATATATTT TTCAACCTTT
47751 CTATATTATA TTTTCAAATA TATCTTCTTG CAGTTAGAAA CGGTATTGAA
47801 AAAGATGTG TGTTGTTCT AGAAAAAAGTA ATAGTAATAT GCCACCAGCA
47851 TTTTATATCA TTCTGCTTT ATTCTTAGGT TCACGGTTCA AAATCAGACA
47901 AAATGAACAT ATTGGTGGC TTTGACAGA TGTTAAAAGA AGGAGGTATC
47951 CGCTCGCTT GGAGGGAAA TGTTACAAAC GTCACTAAAA TTGCTCCTGA
48001 GACAGCTGTT AAATTCTGGG CATATGAACA GGTAAATTGTT ATCACCCGTG
48051 GAATTATTA ACAAAAGAGGA GTTAGTAAAC GGATTCAATA AATGTTAATG
48101 TATAATGCTT TTGGGATTCT TGTTTTAATA CATGATAATC TTTCACATAT
48151 ACCCCATAAG GAGGATCACT TATAGGAGAT TAGACTAAAT AAAATCAGAG
48201 ATTTCTCATG ACCAAGTTAT GGGATTCTTA ATTCTCATATA TTATTTATAA
48251 AGTTTTTTT TTCTAAGTAG TTCTTAAAGG AAGGGTAGAA TTTTAGTTA
48301 TTCATTCTGA ATCCTGAGCA GAAGCAGCAC ACTAACATAA GTTTTATGAA
48351 AGTGTACAA TCTAACCTCT GGAAGGAAA CTATAAGTTG AAGTCCTTTG
48401 TGTAATTGCA CGTTGCTGTA AAATTGAGCT GAGTTGGAG TGACACCTCC
48451 ATGAAGGCAG GGGCGTGGCT TCTTCCCCAT GTACTCCAGC ACCTAGACAG
48501 AGCTTGGCAT GTGATAAGTT TCAAGCGAGT GTTGAATGAG TCAATGAATG
48551 AACAAATGCA TTTACCTCTG AATCACTCT CTGTCGGCTT TTGTTAACTT
48601 GGATTATTTG AGCTATTGCT TCAGCCTAAC TCAATGTAAA GGGGAAATAC
48651 AGAGGTAAGT TTTAGAGTTT GGGTTCTCT TATGGTCATT AGCAGAACTG
48701 TCTAGTTGAG CAGCCACAGA TTATGTTTC CATTATTTAT TCCATCATTG
48751 TTATCAAGG ACTGTAAGGG CCTTGAAATT CAACTCCCC CCCATAGTT
48801 TTGTATTAT TCCATGTAGA TTTTAGATTA TTCTGGAGAG TGTTTTGTT
48851 TTGAGCAACA GAATACTCTT GAGAAGATTA CGAAGTCCAG TGGTATCCTT
48901 TTCTTGCCT AGGAAATAGA GAAGCAAAAA AAAAAAAATTTAA
48951 AGAAAATCTA GTCTCCAGGA TTTTAATTAG AACCTATCCT TGGGAAGGCT
49001 ATTTTCTTA TATGAAGGGT TGAAGATCA AATCATGATT ATTAAGGGCT
49051 AATGTTGAG ATACCCCTAG GTTATTCTGA CCACATACTT GGATTTTATG
49101 ATAGGAAAGC CACAGCCTAA AATAAATAAA TACTCAATGC AGTTATTTC
49151 GTATGCAAGA AGTTGGTAT TTTGAAAAA GTCCATGGGT ATTGCAAGCA
49201 AATATGCACA TTTTGCTTTA TGCCATTGTTG CAGATTCTTA CCTTGGATAC
49251 CACCAACAGG CATCCTCTGC TTCTGTCCAC CCAAGCTCCT TCCTGAGACC
49301 TCTTTATAGT ATTGTGATTT CTGCACACTA ACTTCTTAG ACATGAAGAG
49351 AAAGCTGTCT ACACAGTGTG GTGTAGTTT CTTATGGGCT CTGGACCTAT
49401 GGTGCTGTT TCTCTCCTCC TGCTGAAGGT CCATTCATCC CTCGGGGCTC
49451 TCTAAAGCC ACCTTCCTGT GACAAGCATA TACTAAGCCT ATCAATCAA
49501 GCCAGTTCTT CCCCTGTCCA GCCTCCCTCG AGTGTGAAT TGCAGAATAT
49551 CCCATTTTC ATTGGATGAT GGAAAACCA TTGTTTCCC AGTGGATTGT
49601 AAATTACTTC GGGGTAAATA GGCTGTATAT ATTCTCAAAT TTCCCAGAGT
49651 ATGTAACTAG GTCACTTTA GATTCAAGATA GATTTGTTT CTTGAATAGC
49701 TAGTACTTTA GGAAACTAAG AAAAAGATCT TTCAACCTG GTATGTAGCT
49751 CTGTCAAACA CATCATCAGT ATGGGGTAA CCTGTGTTCT CTGTGGGTTG
49801 TCATTACCAT AGTATGTCAT TTGTATCATT GACAGTGTAA TAGTGTGGGG
49851 TAGTGTCTT GTGGTTCAAG CTGCCACTCT GTACTGACTG CTTCCACTC
49901 CAACATCTTC CTCTTATCT CAACACTGTA GGTCTACCTG TGTACTGTGT
49951 GTTTCAGCAT CTCTGCTTGC ATGACCCAGG AGTGCCTCCC ACTCAATATG
50001 GCCACCATGC ATGGTCATCT TTCTGCTACT CCCTGTCTCC TGACCCGTCT
50051 CCAGCAACAC AGACAGACAC CTTCCCTTT TCTATATGTC ATATGGTGGG
50101 GAATGCCCTT TAGTACTTAC TCAGGAGTTA GTTCCCTCTGG GAAGCCTCT
50151 GTTCTAGTTT CCTTTGTTA CAGCACTTC ACATTGAATT CTGACGTTCT
50201 CTGTACTTAT CTGCTTTGTG AGACTGTGAG CTTCCCTAGG CAGTAGCTAC
50251 TTGTATTCTT AGCACCTTGC CCAGTGCCAG GAAACCCCTTA TTAAGTAAAT
50301 GAAAAGACAG AACTGACAGA CTGGAATTAG AGCTCAAGCT TGCCTCAATC
50351 TCAAGCCATT AAAGATGAAGG GGAGCCGGGC GTGGTGGCTC ACGCCTCTAA

50401 TCCCAGCACT TTAGGAGGTA GTTTGCTTGA GCCCAGGAGT TCAAGACCAAG
50451 CCTGGGCAAC GTGCCAAAC CCCATTTCTA CAAAAAATAT AAAAATTAGT
50501 TGGACGTGGG GGTGTGTGCC TGTACTCAGG ATGCTGAGGT GGGAGGATCA
50551 CTTGAGCTCG AGAGGCAGAG GTTGCACTG A GCTGGGATCA CACCATTGCA
50601 ATCTAGCCTG GGTGATAGAA TGAGACCTG TCTCAAAAAA AAAATAAATA
50651 AATAAATAAA GGGGAAGATA AGGATTGGA ACAGAAGGAG CAGCATGTGG
50701 ACAGAAATGT AGGCACAAGA AGGCATCACT CACTGAAGAG ACTGAAAGTG
50751 GTTCACTGTG CCTCAAGACT GGTGGAGTGT GTTTCGGAA AGATAATGAT
50801 GAAAGAGCTG GACAGATAAA CAGGGGCCAA ATGTAATAGG AGTCTGGATT
50851 TTATTCTGAA TATGGTAGGG GCTATTGTAG CATCTTATAT AGGAAAGTGA
50901 AATGAGTACA TTCACATTAA AGGAATATCA ACCTGAAAAA AGAGTGGAGA
50951 CATTGTTGGG GGAGAGTGGAG GTAGACTAGA GGCAGGGAGA ATATTTAAAT
51001 AATTGAGGTA AGAAATGATG AACACCACTG TAAGGTGATG TCTTTAACAGA
51051 ATGGAGAAGG GAATGAACTG AGAAATATTG TGGAAGTGA ATCAACAGAA
51101 CTCACTGACT GACTGGATAT GGAGGTGAGA AAGAGAAGAG TCAAGAATGA
51151 TATTCTAATT TCTAACTTGA GTGACTGCAT TCAAAGAGAA TACAATATCA
51201 GGTTCCATTG TGTGCATGCT GAGTTTGAGA TGTGTGGAC ATGTACAGGG
51251 AGCTGTCAG TAAAGCAATTG GGTATATCAG CTAGCCATTA AGAGAGAGAT
51301 CTTTGATAGA GAGGTTGTTG CTGAGTTGAG CCATTGGAAT GGGCAGGATC
51351 ACTCAAGAAG AGCTTATAAA TGAGAAGAAT TCTAGGAATA AGTCCAAGG
51401 GAGAAGTAAA AGAAGAAACT TGCAAAGGAC ACTGAGAAGA AATAGCTCGA
51451 GGGATGGGAG AAAATCCAGA GAGAGGGATG GCATAGGAGT CAGTGGAAAGG
51501 AAACGGTTTC ATGGGGGTCA GTACTACTGG GTAGTGAATA TAATAAGAAT
51551 ATCTTTAGG ATTTCTCAAC CCAGAGATAG GTAAGCTTAG TATAAATGCT
51601 TCTGTGAAGT AATGAAATGAA GAAACCATGC TGAAATGAGC TTAAAGTGAA
51651 TGGGAGGTGA AGAAAATTGG ACAGTAGAGA CACATTTTA GGGAGTTTGA
51701 CAGTGAAGAG AAGGAAACTA GAAGAGGGAG AGGGTGATAG ATAAGAAAGA
51751 TGTGTTGGTGG AGGGGATTTG TTTTTTTGTT TTTTTGTTTT TTTTCTGTTT
51801 GTATGTTGTG TTGTTTTGAGA GATGGAGTCT CACTTTATCA CCCAGGCTGG
51851 AGTAAAGTGG TGCAATCTCA TCTCACTGCA ACCTCTGCCT CCTAGGTTCA
51901 AGTGATTCTT CTGCCCTCAC CTCCTGAGTA GTTNNNNNNNN NNNNNNNNNNN
51951 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
52001 NNNNNNNNNNN NNNNNNNNNNN NNNNNNTGCCT CAGCCTCCCG AAATGCTGGG
52051 ATTGCAGGAG TGAGCCCCCCC GTGCCTGGCC TGGAGGGAGG ATTTTGATT
52101 GACTTTAATG TGCCCTGTTG TGAGGAAGC ATGTCATAC AAATAAAGAA
52151 GTTGAACACAA TAGGTAAAGAG AGGTTGATTA ACCCGGTAGG TGTTCAGG
52201 GAGTTTGTGT GTAGGGAAAG GGAGTGGGAG ATGAAAGGG GCTGGGGAG
52251 ACAGGTTCTA TCCAGAGACT GTTAAAAGGA TTAGTCTTG ATTACAAGAA
52301 GAACTCTTCT TATACGTTGTT GGGGAAGAAA AAATATGTA GTAGCTATGG
52351 ATAATTCTTG AGGAGGTGGG CAGAATACCA AGATATTCTG CCTGGTGGCC
52401 TCTCTACTCT TCCCTGAGCT CCTGAGAAAG GATGTGATCT GAGAATGAGG
52451 GAGGAAGTGG TATTGGAAGC TGGAGGGAGA TGGAGGAAGAT CAAAATGGT
52501 AGTCTAACAA ATGGGAGAGA ACTGAGATAG ACAAAGGAT TTCAGGGTGG
52551 TTTTGAGGGC TCAGTTAAGT CTCCTTCTAGG AAGGTTCACT TCTGTAGCCT
52601 TGGCAAGTTA CTTAAAGTCT CTGTGACTAT TACCTCATCT CTAAGATGGG
52651 GACTAAGCTT GGTGACATAG TTTTACATAC CAGGCACAGT GCCTGACTTT
52701 TTGGCTCTGT CCTGAAGTCT TCCCTTGTG TATGGTATGT TTCGGGGAAAT
52751 AGGAGCCTCA AGCACCTTAC CTTTAAATAT TTATCCTCCA TCAGTCATA
52801 AACGTTTACT CTGTTACTTT GATAGGTGCT GTGGGGGTCC AGGGTATAAA
52851 AGGTACCTTC AAAGTTACTG TTAAAGTGCA GGAAGGTTTT TAAGCAAATT
52901 ATGTTAATG ATTTGACAA TCTGACATGC AGGAAAATTA ATAGGGCCTA
52951 TGCAGAAGAG GAGTTTATG TAACACTCTG TAGTTCAAGGA AACAGAGCCC
53001 TTGGGAAGCAG TGATCTCTCT GGGGAGGAAT GTCTGGTATT TGGGAATCTC
53051 ATGAAATGAT AATATACTTA ATTTTATCA TGAGCAGCAA AACACAGATT
53101 TGCTAGGAGA AAGTCATCGT ATGTTGTGC ATTGGGCACT TTAGATCCCA
53151 GGGAACAGAA ACTGGCTGGC ACAGGAATGG GCATCACTGT GGGGATGGAT
53201 CATGTAGGGG AAGGATCCCT GGAGAAGTCC AGGAGGTGAG ACTTCCCCCT
53251 TCCCTCTCC ATGCATGAGT CCACCTCTCT CTGTTGACTT TCCCCTTGTG
53301 CCTCTGGTGA CAGCAGCTGC TTACCTCTGG AGACCCCTC ACATTTCTGA
53351 GAGAAGGAAT CTGGCTTGCC TGGCTAATTG CCATGGTCTA TGTTGGGCA
53401 GAATGCTTA GCAAGTTGTG TAAAGATAGT GTATTCACTAT ATTAATAATA
53451 ATAATAACAT CTACTGAACAA TTGCTAGGT GTTCAGACCT GCACTAACCG
53501 TGTTACAAGT ATTATTTTT TGTAATCCTT TCCATAACCC TGTGAGGTTAA

bioRxiv preprint doi: <https://doi.org/10.1101/2023.06.06.540000>; this version posted June 6, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a [aCC-BY-ND 4.0 International license](https://creativecommons.org/licenses/by-nd/4.0/).

53551 GTACTGTTAT CACAGACAAG GAAACCACAA TGTGGACCTG TTCATGAAC
53601 TGCTCGAGGC CACGTGGCTC TGGAGTTCCA GCTCAGGTCT GCCTGACTCT
53651 CAATCCCAGT ATATTAATAT ACTGGCCAGT CACTATTTG GCTGTATTGG
53701 GGTCAATTT ATACCCTTGG TCCAGTTAGC TATGTTGGGT CACTTTAGTA
53751 CTGATAGCCA GGGAGATGCT GGGCTTGATA GGTTAGTATA ATTCTATGTA
53801 TTACCTACAA AAACGTGTTT TATAAATTGT TTTGTTAACAA TTTGTTTGTG
53851 ACCTATTAT TCATTTTATT TGCACGTGGT AAAATAAACT CATCTTTAA
53901 AAACTGTGGG GAAAATATCC AAACATTGTG AAAACTTGAT TAACCTTGTA
53951 TTTCTGTAC ACCTGGGGAG GGATGCTGTT ATGCTGTTTC AGCAAAGGAG
54001 CAACTTGGTC CAATCTGGGA GACATCTGTG TTTTGTGGAA ATCTGACTTG
54051 AAAACCACTG TCCAGTCACT GCGTGTATTA GCATTAGGC CTTGCTCTTC
54101 TGCTATGTAT TATTAATGTA GTGTATACAT TTGAGACAC ATCATCACAT
54151 TTGTCATTT ATTGATTCT AGGAGCTGTT TTGTTATTCTA GGATTGTCTA
54201 GTTGGCTTGG GCTGCCATAA AATACCACAG TGTGTGTGGA ATCAACAAACG
54251 GAAATTATT TCTAACAGTT TCAGAGGCGG GAAAGCCTAA GATCAAGGGC
54301 CAAGCCAGT TGATTTCTAG TGAGCGTTCT CTTCTCAGCT TGTAGACAGC
54351 TGGTATGTGC TCACATGGTC TTTTCTTGGT GCACATGTGA AGGGGGGAGAG
54401 AGAGAGTGGG CTCTCTGGTG TCTGCTCTTA CAAGAACACT GATCCTGTCA
54451 TGAGGGCTCC ATCCTCATGA CCTCATAACC CTAATTACCT CCAGAACGCCT
54501 CATCTCTAA TACCATCACA TGGGAGGTTA CAGCTTCAAC ATATGAATT
54551 GGTGGGGGTG CAGCTCAGTC CACAGCAGGT AGTAATGTGC ATTTTAAAAC
54601 TTGTTTATAC AGTACAAGAA GTTACTTACT GAAGAACAGC AAAAAAATAGG
54651 AACATTGAG AGATTTATT CTGGTTCCAT GGCTGGAGCA ACTGCACAGA
54701 CTTTTATATA TCCAATGGAG GTGAGTACCA TTGTCAGTC TGACTGTGTG
54751 ATGGTGTTCG TGGTGGTTGT CTATTGCTCT CTAACAAGTT ATCCCCAAAT
54801 TAACAGTTTA AAACAAGCAT TTATCATCGC ACAGTTCTC TGGGTCAAGGA
54851 ATCTGGAAGC AGCTTAGCTG GGTGCCTCTG GCTCAGGGTT TTTCACAGCC
54901 CACAGTCAG ATGGTAGTC GAGCTTGGAA TCAGCTGGAG GCGGATTCCA
54951 AGCTCACTCA TGGTGTGCC AGGCCTCACT GGCTATTGGC TGGAAACATC
55001 AGTTCTTAT CACGTGAGCC TTTCTGTAGG CTGCTGTAGT ATCCTAAAA
55051 CACAGTAGCT GGCTTCCCTA GAGTCAGTGG TCCAACAGAG AGAGAGAGAG
55101 AGAGTGCCTA AGATGAAAGC TGGTATCTT TGCCTCTTCT GCTGTATTCC
55151 ATTGATCACA CAGACCAACC CTGGTAGAGT GTAGGAGGGG CTGGTATAAT
55201 GGTGTTAATA ACCGGAGACA AATATCACTG GGGGTCACTT TAGAGGCTGG
55251 CTGCCACTTT AGAGGCTGGC TGCCATTCT GTCCAAAGAG TTTCTGTACC
55301 ATAAATTTAA TAATGGAATC TCAGGATTTG ATTATATGGT GATTATCCTA
55351 ATTAGACATC CTTTCATTAG TGCATAGGTT GGCAAAACAC AGACCTACGG
55401 ACTGTTTCAT ACAGCCCTTG ACCTAAGAAT GCCTTTACA TTTTTAAAAA
55451 GTGGGCAACA CAGGAAAAAG TGAGAAAGAT CTAAATCGA CACCTAAGA
55501 TCACAAATTAA AAGAACTAGA GAAGCAAGAG CAAACAAATT CAAAGATAG
55551 CGGAAGACAA GAAGTAGCTA AGGTCAGAGC AGAACTGAAG GAGATAGAGA
55601 CACGAAAAAC CCTCCAAAAA ATCATTGAAT CCAGGAGCTG TTTTTATGAA
55651 AAGTTAACA AAATAGACAA CTAGCCAGAA TAATAAGAA GAAACCAAGAG
55701 GAGATCAAA TAGCCCCAAT AAAAAATGAT AAAGGGATA TCACCCACCAA
55751 TCCCACAGAA ATACAAACTA CCATCAGGGA ATACTATAAA CACCTCTATG
55801 CAAATAAAACT AGAAAATCTA GAAGAAATGG ATAAATTCTC GGACACATAC
55851 ACGCTCCCAA GACTAAATCA GGAAGAAGCT GAATCCCTGT ATAGACCAAT
55901 AACATGTTCT GAAATTGAGG CAGTAATTAA TAGCCTACCA ACCAAAAAAA
55951 ACCCAGGACC AGACAGATTC ATAGCCGAAT TCTACCAAGAG GTACAAAGAG
56001 GAGCTGATGC CATTCTTCT GAAATTATTC AAACAATAGA AAAAGAGAGA
56051 TTCCCTCCCTA ACTCATTCTA TGAGGGCAGC ATCATTCTGA TACTAAAACC
56101 TGGCAGAGAC ACAACCAAAA TAGAAAATT CAGGCCAATA TCCCTGATGA
56151 ACATCAATGT GAAAATCTC AATAAAATAC TGGCAAACCTG AATGCAGCAG
56201 GACATCCAAA AGTTTATCCA CCATGATCAA GTGGCTTCA TCCCTGGGAT
56251 GCAAGGCTGT TCAACATATG CAAATCAATA TAACGGAATT CATCAATAAA
56301 CAGAACAGT GACAACAAAC GCATGATTAT CTCACATAGAT GCAGAAAAGG
56351 CCTTCGATAA AATTCAACAC CACTTCATGT TAAAAACTCT CACTAAACTA
56401 GTTATTGATG GAATGTATAA CAAAATAATA AGAGCTGTTT ATGACAAACC
56451 CACAGCCAAT ATCATACTGA ATGGGCAAAA GCTGGAAGCA TTCCCTTGA
56501 AAACCGGCAC AAGACAAGGA TGTCCCTGT CAGCACTCCT ATTCAACGTA
56551 GTATTGGAAG TTCTGGCCAA GGCAATCAGG CAGGAGAAAG AAATAAAGCG
56601 TATTCAGATA GGAAAAGAGG AAGTCAAATT GTCTCTGTTT GCAGTTGACA
56651 TGATTGTATA TTTAGAAAAC CTCCCTGTCT CAGCCCCAAA TCTCCTTAAG

56701 CTGATAAGCA ACTTAAAGCA AAGTCTCAGG GTACAAAATC AATGTGCAA
56751 AATCACTAGC ATTCCCTATTA ACCAATAATA CACAAACAGA GAGCCAATC
56801 ACGAGTGAAC TCCCATCCAC AATTGCTACA AAGAGAATAA AATACCTCG
56851 AATACAACCT ACAAGGGATG TGAAGGACCT GTTCAAGGAG AACTACAAAC
56901 CACTCCTCAA GGAAATAAGA GAGGACACAA ACAAATGGAA AAACATTTCA
56951 TGCTCATGGA TAGGAAGAAT CAATATCATA TCATAGGAAG AATCAGTGGC
57001 CATACTGCC AAAGTAATTT ATAGATTCAA TGATATCCCC ATCAAGCTAA
57051 CATTGAATTT CTTCACAGAA ATAGAAAAAA CTACCTTAAA TTTCATATGA
57101 AACTAAAAAA GAGCCTGTAT AGCCAAGACA ATCCTAAGCA AAATGAACGA
57151 AGCTGGAGGC ATCACGCTAC CTGACTTCAA ACATACTACA AGGCTACAGT
57201 ACCAAAACA GCATGGTACT GGTACCAAAC AGATATATAG ACCAATGGAA
57251 CAGAACAGAG GCCTCAGAAA TAACACCACA CGTCTACAAC CATCTGATCT
57301 TTGACAAAAA CAAGCAATGG GGAAAGGATT CCTTATTTAA TGTATGGTGT
57351 TGGGAAAATC GGCTAGGCAT ATGCAGAAAAA CTGAAACTGG ACCCCCTTCCT
57401 TACACTTAT AAAAATTTAA TTAACTCAAG ATAGATTTAA GTCTTAAACA
57451 TAGACTTAAATC CTATAAAATC CCTAGAAAAA AACCGAGGCA ATACCATTCA
57501 GGACACAGGC ATGGACAAAG ACTTCATGAC TGAATCACAA AAGCAATGGC
57551 AACAAAAGCC AAAATTGACA ATGGGATCT AATTAAACTA AAGATCTTCT
57601 GCACAGCAA AGAAAATATC ATCAGAGTGA ACCGGCAACC TACAGAATGG
57651 GAGAAAAATT TTGCAATCTA TCCATCTGAC AAAGGGCTAA TATCCAGAAT
57701 CTATAAGGAA CTTAAGCAAA TTTACAAGAA AAAAATACCC ACCAAAAAGT
57751 GGGTGACGGG TATGAACAGA CACTTCTCAT AAGAAGACAT TTATGCAGCC
57801 AACAAACGTG AGAAAAGGCT CATCATCCCT GGTTGTTAGA GAAATGCAA
57851 TCAAAACCCC AATGGCATAAC CATCTCACGC CAGTTAGTTA AAAAGTCAGG
57901 AAACAAACAGA TGCTGGCAAAT TATGTGGAGA AATAGGAATG CTTTACACT
57951 GTTGGTGGGA GTGTAATTTA GTTCAAGCAT TGTGGAAGAC AGTGTGGCAA
58001 TTCCTCAAGG ATCTAGAACC AGAAATACCG TTTGACCCAG CAATCCCATT
58051 GCTGGTTATA TACTCAAAGG ATTATAGATT TTTCTACTAT AAAGACACAT
58101 GCACACGTAT ATTATTATGCA GCACTGTTCA CAATAGCAAA GACTTGGAAC
58151 CAACCCAAAT GCCCATCAGT GATAGACTAG ATAACACAAA TATGGCACAT
58201 ATACACCATG GAATACTATG CAGCCATAAA CAAGGATGAG TTCATGTCCT
58251 TTGTAGGGAC ATGGATGAAG CTGGAAGCCA TCATCTCAG CAACCTAACAA
58301 CAGGAACAGA AAACCAAACA CCACATGTTCA TCACCTCATAA GTTGGAGTTG
58351 AACAAATGAGA ATACATGGAC ACAGGGAGGG GAACATCACA CACTGGGGCC
58401 TTTTTGGGG TGAGGGGCTA GGGGAGGAAT AGCATTAGAA GAAATACCTA
58451 ATGTAGGTGA CAGGTTGATG GGTGAGCAGA ACCACCATGG CACGTGTATA
58501 CCTATGTAAC AAACCTGCAC GTTCTGACCA TGTATCCCAG AACTTAAAGT
58551 ACAATTTTA AAAAGTAGGC AAAAACAAAAA GAAAAGAAAA GTAATATACA
58601 ACCGAGACCT AATATTTTAG GCTTGCAACG ACAGATATT TACTATTAG
58651 TCTTACAGG AAAAGTTTC CAACTACTGC TTTATAGCAA AAATAATATT
58701 GTAGATGTGG AATTATTGAA TATAGCAGAG GGGTTTTAG TAACTGATGA
58751 CTTAACGAAAG ATAACACAA TTTTCACCGA TATGTGGTAT GCATGCTAAT
58801 ACAGCTTTT TTAAGCATCT TAATATGATT GTTTATATTAA CTCCACACAC
58851 CTCTAAAAAA AACTTAATAC CCTATTTTC CTCTCATATC CTCCCATATC
58901 AGTTAATAGT ATCACCTTCC CAACTCCCCA CTGCCCCATC CTGTGTTCCA
58951 AGCTAGAAAGT ATTGGGGTTA TCCTTATAC TACCATTCC CTCACCTTCC
59001 AGATGCAGGT GGTCAACAGT CAGTTTGTT AAGACATCAA TAGATTATCT
59051 TGCTTCCATT TCCTGGTCA CTTCCTTCAT CAGATCCTCC TTGCAAGTAA
59101 CGGGTCTCTC TGGCTTTGGT CTTAGCCCCA CAATAGAGGT AATACATGAA
59151 AGAGAAATGTA TCAACAAATT GTACAGTCTT TTGAGTGCATA ATATGTGCTA
59201 GGTATTGTT CCATGTAAAA TTACTTCATT TGAATCCCATT GATGATAGAG
59251 TTAATATGAA CAATCATATT TTGTTTTTT TTATATCCAG GTTATGAAAAA
59301 CCAGGCTGGC TGTAGGCAGA ACTGGGCAGT ACTCTGGAAT ATATGATTGT
59351 GCCAAGAAGA TTTGAAACAA TGAAGGCTTG GGAGCTTTT ACAAAAGGCTA
59401 TGTTCCCAAT TTATTAGGTA TCATACCTTA TGCAAGGCATA GATCTTGTG
59451 TGTATGAGGT GAGTTTGAG AATCTTTTG AATTGGAAAAA TGCAAGTTAGA
59501 TCTTGTAGA ATTGGACTTT ATATGAAGAA GTAGATATAT ACCAGAAAAC
59551 AGTGTGTGAC CAGAAGTAAAT TTCAAGCAGT TGTTATTGAA ACTTTCAAGT
59601 AACTTGAGTG TGAATATGCA TGGGGTCACT TTTGTATTAG ATTTTCTTGG
59651 GAATTGCTTT TTGTTAATGAA GAGTAGACCTC AAAGGTAGGT ATAGTTGTTC
59701 ACCTTAAAG GTGTTTCTAG AGATTTTTTC CTTTGTGTTTG GATTGCAA
59751 AATCTGACAT TAAGCCAAGT GACTAATGTG ACTAACATGA GTAATACAGT
59801 TTCATTCCCT GTACGGAAGA ATACAAATCT TGGATCAACC CTGCAATCTA

59851 AATCATTAA TAATTTATGA ATCTCACAAA CAATTATTGA GCACACACTA
59901 TACAAACCAC TAGGTTAGAC ACTGGATCTG GGGATCAAA GGACTCAATG
59951 TGTGCCTTGA AGAAACTGAA GGTCTGGTGG GGGAGACAAA CGACTAAAAC
60001 TCAGCGTGGT TATCTGTGCT GCGACAGACA TGAGCCAGGG TGCATGTTAG
60051 GATGAGACCT AAGCTACAGC GTAGAGGAAG AGTGGAAATGT GTAAATGAAAAA
60101 GAAGAGTCGA ATTNTTTTT TAAAGAGCTT TATTGAGATT TAGTTCATAT
60151 TCCTTACATT TCACTCATT GAAGTGTACA AGCAAATGGT TTTGGCTTC
60201 TTACATAATT TTTAAAATT ATTATAAAAT ATAAAATTG CCATTTACT
60251 AATTTAACGT GTACAAATTCA GTGGCATTAA TTACATTCA AATATTGTGC
60301 AACCATCAAC ACTATTTCCA ATCCCTTTC CTCACTCCAA ACAGAAACAC
60351 CTAAACCTT AAGCAATAAC TTCCCTACCT CCGTAACCTCA AACCTTGGT
60401 AACCTCTAAT CTGCTTCTA TGTCTAGGAA TTTACCCATT CAAGATATCT
60451 TATAAGTAGA ATCATACTAGT ATTNTTCTTT TTGTGTCTGA TTTATTACTC
60501 TTAGCATAAT GTCTCTAAGG TTTGTTCATG TTGTAGCATG TATCAGAACT
60551 TCATTTCTT TCATGGCTGA GTAATATTCC GTTATGTGTA TATACCACAT
60601 TTTGTTAGT CCTTCATCTG TTGAAGAGCA TTTGGATTAT TTCTACTTTT
60651 CCAACATTGT GAATAATGCT GCACTGAACA TTGGCATTCTG CGTATCTGTT
60701 CGAGTCTATG CCTTCATTTC CTTTGGGTAT ATATCTCAGA ATGGAATTGC
60751 TGAGCCATAT GGTCAATTCTG TGTAGCTT TTAGGAACTA TGAGACTGTT
60801 TTCCATAGTG GCTGCACTTA CATTCTCACC AGCAACATAC AAAGGTTCCA
60851 GTTTTCCAC GTCCATTATA ACACCTAATT TCCATTAA AAAAGCTTAT
60901 TTTTATTATG GCCGTCCTCT TAGGTGTGAG GTGGTATGGT TCAGGACTTT
60951 ACTTCTTGTG CTGAGTTTTT TAAAAAAATTG TGATTAACACAT CACATAACAT
61001 AAAGTTTATG ATTTAACCA TTTTTAAATA TATAGTACAG TAAGTGTAA
61051 CTGTTGTGG TTTGTTGTGC AACAGATCTC TAGAACTTTT TCACTTCTCA
61101 AAACCTAACAC TCTATAGTCA TTAAACAAACA GCTCCCAATT TCCCCTTCAC
61151 CCCAGCGCTG TGTAACCTAC TTTCTCGTT TATGAGTTG ACTACATTAA
61201 ATACCTGTGTA TAAGTGAAT CATGTGGTAT TTCTCTTTCC GTGACTGGCT
61251 TATTTCATGT AACATAGTTT CCTCATGATT CATCCATATG ATAGCATAACA
61301 ACAGGACTTT TTTGTTTTA AGGCTGAATA ATAATTGTT GGGTATATAT
61351 ATCACATTTC CTTTATTCTAT CTGTTGATGG ACATTTGGAT TGTCTACA
61401 TCTTGACTAT TGTGAATAGT GCTGCAGTGA ACATGGTTGT GCAAATATCT
61451 CTTCAAGATA CTGTTTCAG TTCTTTTGA CATATACTCA GAAGTGGAAAT
61501 TTCTGGTCA AATGGTAATT CTATTTAA GTTTTGAGG AACCTCCATG
61551 TCATTTCCA TAGTAATAG ACCTTTTGT TTTTTAACAT TTCTATCAAT
61601 GTACACCAAG ATTCACATT CTCCATGTCC TCCCCAACAC CATTAAGTGG
61651 GGTGGTGGTC TACTACTT GCTGTGTGTC TGTATTCC TCCCTTCAGT
61701 TCTGTAAGTG TTGCTTCAAT ATATTTAGGA GCTTAATATT AGGTCCATAT
61751 GAAGTTATAA TTCTCTCTG GTAAAGTGAC CCATTATCA TTATGTAATG
61801 TCCATCTTG TCTCTGTGA CAGTTGTTG CTTAAAATCT ATTTGTCTG
61851 ATGTAATTAT GGCCACCCCT TTTCTTTG GTTCCCCGTT TTTATGGAAT
61901 ATCTTTTCC ATCCTTTCACT TTTCAGCTTA TGTGTGTCT TAGATCTAA
61951 GTGAGTCTCA TAGATAAGGT ATAGTTGATT CTGTATGTGT TATTCACTCA
62001 GCAATTATA TCTTTAGTT AGGGGATTTA ATCCATTAC ATTAAGCA
62051 GTTACTGATA GGGAAAGACT TACTGTTGTC ATTTGGCTAG CTACCTTTT
62101 ATCTTGTCC TGTGGCTTTT CTGTTTTCC CTTCTCTCT TCTGGCTTC
62151 TTCTGTGTT TGTGATTTT TTTTTTTT GTAGTGTAT GTTCTGATTC
62201 CCTTCTCATT TCCCTTGTG TGCATTCTAT AGATGCTATT TTTGTGGTTA
62251 CCATTGCAAC TACATAAAAGC ATACTAAAGT TATAGCAACT TATTTTAAGC
62301 TGTTTACAAC TTAACCTCAG TGGTATATAA AACTCTATTT CTTTACATAT
62351 TTCACCTCCT CCCCACAAAC TTTATGTCTT TTGATATTGT ATATCCTTAA
62401 CATAGATTAA TAGTTACTTT TTATGCTTT CTTCTTTAAA TTCTGTTTAA
62451 ATTTGTGTT TGAAATTAG ATTTTCAAGT TATTTATATA CCTTCATTAC
62501 AATACTATAG GATTTATAA TATTCTAAAT ATTGACCTTT ACCATAGAGT
62551 TTCATATTTC GTGGTTTTGT GTTGCTATT ATCATCCTT TGTTCTCCT
62601 TTTAGCCTTT CTTGTAGGGC CGGTCTAGTG GTGATAAGCT GTATCAGCTT
62651 TTGTTGTCA GGGACAGTCT TAATTTCTCC TTTTTGAAAG GGCAGTTTG
62701 CCCATACAGT ATTTTTGTGTT GGCAGTTTT TTAAGTTCA AAACATAGAA
62751 TATAACATTC CATTTCCTTC TAACCTGCAA GATTTCCATT GAGAAATGCA
62801 CTCAATGGAT TTTTTAAATCC ATTGAGATAA TTTTTTAATC CTGAGGATT
62851 TAAAATTTT AGTCTTACAG GATTTAAAAA TTAAAAAAGTT AAACCTGTTA
62901 TATAACATAT TAACATGTAT TTTTAACTTA AAGTATCTTA TGTTAAAAAA
62951 GTTGATTATC ATATATATT TATACAGTTT CTCCTAATTA TTGCCTTCTA

63001 ATGAAATACA GGGACCTAGA GTAACAGGGA TAAAGTATGG CCTTTTGATC
 63051 AGCACGCCCTG GTTCTGAGTC CTTCTTAAAA AAACCTCTGGG CCTGGTGTGG
 63101 TGGCTCATGC CTATAATCTC AGCACTTTGG GAGGCCGAGG CGGGCGGATC
 63151 ACCTGAGGTC AGGAGTTTGA GATCAGCCTT GCCAGCATGG TGAAACCCCTG
 63201 TCTCTACTAA CAGTACAAAG ATTAGCTGGG CGTGGTGGT GGTGCCTGTA
 63251 ATCCAAGCTA CTCAGGAGGC TGAGGCAGAA GAATCGTTG AACCTGGGAG
 63301 GCAGAGATTG GGGCACTGCA CTACAGCCTG GGTGACAAGA GCGAGACTCC
 63351 ATCTCAAAA ACAAAACAAA AACTCCGCTG AGATGAATT TTCTCATTTC
 63401 TAAAATCAGA ATAATAGATT TATGTAAGAG TTTCTGTAAG GCTCAAATGA
 63451 AATATATGTA ACCTGTAAAA TGAGATACAA TTAGTAGAAT TATATTATT
 63501 TATTAATACT CACCATAAGA GGTGTTCTT AGATCCTGCA GCGTTGCTG
 63551 CGCAGTTCAC GTTTGTTTAG AAGAATGTCA GTAACCCTG CAAACCTCAT
 63601 GTGTTCCGCA CCCCCAGTGG CCTCCCACCT CTCCACAGAG TCACCGCCTC
 63651 CTGCAGTGCC TGCTGCTTCT GCAAATGCGT GGCTCATCC TGCAGAAACG
 63701 GGGCTCTCA TGAGGTTGAG AATAGCTGTG AAAATGTTA CGTTGAAGTT
 63751 GTAGAGTTCG TTAATTATTT TCTTCTTTAT TTCTCTGGCA GCTCTTGAAG
 63801 TCCTATTGGC TGGATAATT TGCAAAAGAT TCTGTAAACC CTGGAGTCAT
 63851 GGTGTTGCTG GGATGCGGTG CCTTATCCAG CACCTGTGGT CAGCTGGCCA
 63901 GCTACCCATT GGCTTGGTG AGAAACTCGCA TGCAGGCTCA AGGTGAATT
 63951 TTGATTACAG AACCACACCG ATAAAAGTGC TGCACCAGTA ATGTGTTTTT
 64001 AGAACTCCAA GTTCTACTAA GATGCAGACT GTAGTTTAA GACAGTATT
 64051 CTCAACCTT TTTTCATTAT TGCCCTCTTA AGGAATCTT TCAGAAATTC
 64101 TTTTCTAAA TGCTCCCTCG TCATGAAATT TTAATGCGAC AGAACGATTG
 64151 CATATGTAAT GTATGCATAC ATATGCTTA TAGATAAAACA GAGTACTATT
 64201 TTTTTGACT GTGTTACATG CAGTTTAA GATTATAAGC TTTAGTATCT
 64251 GATGGATTG GTTTCAGATC CTTGCTCTAG ACTTCTTGGG GTTTTTAATG
 64301 GGAATGAAAA TTGTACAGTG TTGTAAGAAT TACCAACAAT ATAAATAAAG
 64351 CATCTGGGT TTGTTAAATT TTGGTAAAT GGTGGTTGGA ATCATTTTTT
 64401 AGTGTGCGT AGACCCCTACA AGTTTGAGC TGTGATTCCCT CCTCACTGTG
 64451 ACACTGTCTC CATGTTGGC TTTGATTACA CTGTACCATC CTGGTTGTT
 64501 TGCCAGCCCA TTGATAACTT TTACCATTTG CTGGTTTTA TTGCTATCCC
 64551 CACTCTTATA AAGTATGCAT TCAAATGCCT TTCTTTCTC TTGATGCTT
 64601 TCCCTGGTCA GTCTTATCCA TTGTTTCTT AAGTAGTACA CCTTGGCAT
 64651 CTACAGCTCT ATTCCCAACC TCCCTTCCAA GTGCCAGCCA CAGCAACCCC
 64701 AGCCAAGCAG TCAGTAACTA ATTGGCAAAT ACTCCCTGAG CCATTGTCCC
 64751 ATTCTAGACA CTGCCAGATG CTAGGGTAG AGCAGTCAAC AACTCAGGTG
 64801 TGGCCCCGCC AGTGTAGAGT AGAGAACAGC TTATGTCAG CAAGTAAACA
 64851 ACCTGGTTAA ACCAACTCTT CTTTGTAG GGGAGCACAG AGCAAGGAGC
 64901 TATAACTAA CTTGGCCGCT GCAGAATGCT GTCACTGAAG CTGAGACTGG
 64951 AAAGATGAGT GGGAGTTAGC TGGGCACAGG CCAGTGGAGT GGGAACAGAA
 65001 AACATTCAG TTGAGGGAAA GCATGTGTA AGACACTGAG GCAGGCACCA
 65051 ACATGGTGTAA TTTAGGAGC TGAGAGACAG TCATGGCTGT AGAGAAAAAC
 65101 ACAAAAGTAGT GAACTACACG TTTCTTGTGT ATTCTCTCAT TTCAACATCA
 65151 TAACCATCTT GGGGATGGGA ATACTAACAT TATCCCATT TTTCAGATGA
 65201 GCAACTGGGG CAGAGAGAAT TTAAGTAACCT CCCACAAAGAT TATACCTGTG
 65251 GTAAATAGTG GGACTGAAAT TCAGACACAT GCAGTCTGAT TCTAACCCCTC
 65301 CTGCTGCCA GCTCTGATCC AGAACCTTG GACTGTGATA CGGCTGATAG
 65351 ATTGTCTATG GCTGATAGAC TGTCAATTCT GACCTAAAAG TCTGATCATT
 65401 TTACATCTGT TCAGACATCT TTGCAAGCTT TCGGTGTCAG TTCCAAAGTT
 65451 GTTAGTGGGA ATTCAAAGC CTTTAATAAT CTAGCCCCAC TTTGTTCACT
 65501 CTCTGTGTAA TAACCACATA CAACAAATTGG CTGCATCTCC ATAGCACATG
 65551 GTACTCCTCC CGTGTCTTG GTTGTGCCAG CAACACTGGT TTTCGCTTT
 65601 TCTTCCTGCT TGTGAGGTC ATTCCAAAGG CCCAGGTCTT TGTGCTTTT
 65651 CCCAAGCTTC CCAGAGCTTC TTCCATACCTC CCCTTACTTC CTGAGATTTA
 65701 ACTGTTCTCT CTTCAGCGCT TGTCTAGTAA GAAGGAGGCA GCAGCAGCAC
 65751 TGTGGGGTGG TGGAAAGTGT ACCAGCTTG GAGTCAGACC ATTGGATCTC
 65801 AGCCCTACCA TTTTCTACTT AGATTTTTT AGGACAAATT TCTCCATCTT
 65851 TCTAAGCCTC CAATTGCTCA CTTACAAAT TGATATAACA TTACCTTGC
 65901 AAGATTGGTA TGGAAGGTAA TTAACCCAGT ATTTAGAACAA TAGTAATTAA
 65951 TAAATAACTA TTATTACCAT CATTACTATA GTTAGGACAC TCACTGTTAG
 66001 GTGCTATACA AAGAGGATCA TAAAAGGGAT GTTGTCTTGG GCTTCTTGG
 66051 ATAAATGTTG TCCTTTACT GTATTTAGA ATATCATTCT GGGTCATAAT
 66101 TGTTTGTGT CATAATAATG AAACATACTT GAATATTAAA TTACCCCTCTT

FIGURE 3, page 21 of 42

66151 TTTTTATTT TTAGCCATGT TAGAAGGTTT CCCACAGCTG AATATGGTTG
66201 GCCTCTTCG AC GAATTATT TCCAAAGAAG GAATACCAGG ACTTTACAGA
66251 GGCATCACCC CAAACTTCAT GAAGGTGCTC CCTGCTGTAG GCATCAGTTA
66301 TGTGGTTAT GAAAATATGA AGCAAACCTT AGGAGTAACC CAGAAATGAT
66351 GTTGCATTT TTGCTTTAGC CTGATAATTG AAACCTCAA CAATCTCTGG
66401 AGTGACTTT TCTCCTCGAA TTGAAACAAAG TCTATGGCAA AAGAAGCTGC
66451 ATTTTTTC ACAAAGGGAA GATGGTAACA ATGGTCACCTT CAAACCTTTG
66501 GGCTAAATTA TATGTACACA GAAATGTTCA AAATCATAGT TTTAATGTGT
66551 TTTGAAAAGG CCACACAAATT ATACTTTATC TTTTCTTAAT AATCCTGCAA
66601 ATCTCTGCC C TGAAATCCGAA ATCTGAAAAT GTACTGGCTT GAACAAAATT
66651 TGTTTGTGT GTTAGAGTTA TAAATCATT ATCTTTATTT CGGGTGGTTT
66701 ACGTTTATGC CAGTTCCCTT ATATTTAAAT TTCTGTTTT ATATATTTG
66751 AATGTCCTTA TAGATTTCTT TAAATTCCT TATAGAACCA TTAATAGAAA
66801 ATCATTACAT TTTAAATATA CTTACAGCA AAAGCATCCA ATAAGTATA
66851 GGGTTTATGT CCTTATTTT CTTTCAGCTG AATACGAATG AGCACAGTGG
66901 TGGAATTCTC GAAGGGAAAGT GATGAAATTA TATTATTTT AGTGGGCACT
66951 TTTCCATTTC ACCACTGTAC CATTATTG TGCTCGGAGT TATACACTAA
67001 TTTTCAGTAT ATTACTGT TA AATTACCAAC ACAAGGCAAT TTATTTGAAA
67051 GATTCGGTTT ATCCCTGCCAT TGCTTGAAA AGCAGCAGGA AACGAAATCC
67101 TTTGACTTGT ATCAGCTTCT GCAGAGCATIC TTTGTTTCC TTTGTCCTTT
67151 GTTTCCTACC TTTGAATCA GATTCCGTT TAGTCAGGAA GACTTCTTGG
67201 GACCATTCTT AGTAACCTGA AATTCTTT TTAATTGCAT GAAGTGGATT
67251 GATCATGAGC AAATGATGTG CTTATTCCTC CCTCACTGTT GAATATCTT
67301 GAACTGCTG TTTCAATAT GGGCAGCACA AAGGTGAGAG ATACATATTA
67351 ATAGTAGTAT GTATTACTCT TATACATTAG ATACCTATAT TAAATGAAA
67401 GGCCCAATTG GTAAACATAT ACATTCAAT TCTCTCTTGC CCCAAGTTT
67451 AGGAACATGT TAGGATATAG GAGACTTAAT TTATAATAAT GAGAGCATT
67501 TTTTATTTTA C TAAAGCCAT TTTTATAGTC AACTATCTT TCTTATTTGT
67551 GTGATTAGAA CTTAGAAAAA TATTTACTAG TTGAAGTTAT TATCAGTTT
67601 TAATTAGTT CTTAAACTCA TTTCACTCT AATAATTCT GTTATAAATT
67651 GCCAGCATT TAAATGAAAAT CTAATGATGT AATAGGCATT TTCTTTATT
67701 GAACTTACCT CTTTATTTT CTGAACCAAA GAGAAAGATG GACTGGTGT
67751 TGTGAAACAT TT TAAATTTAT GTAGTTCT TATATTAGT TATGTTTGAT
67801 AAATGTCCTA GTATTTTAT AATATGATAA GCCTGGGATT CTACTTTAG
67851 GGTATTGTG ACTTTTGAGT AATATATAAA GTGACAATAT TAAGGTACAT
67901 GATCAGCTCT TTCTATTTT ACTCGAAAA ATTATGGAAA TGAATAATT
67951 TGCTAACAC TTTGAAATT CAAACTTCTG GAAAATATGA AAATATTCT
68001 TGTTCAATT GAAATTTAAAT TGTAAGGTAT GAATGTGATT TGTCTGTACA
68051 TCTTGATCT TTCCAAAAA ATGATTCTGT ATCTTTGGA AAAAGCCGA
68101 GAGTTGAAGA TAGTATATT TGTTCTAAAT TACTGTTTT CCAAGTTTA
68151 CTATCAAAAAA TATATATTG TTCTCTAAAT TACTGTTTT CCAAGTTTA
68201 TTTTTTTTAG AGAAAATTCT TAAGTCTCAG TTTCTTAATT GAAAAAAA
68251 AATTATAAAAT AAAGCAAAA TTGTATCTA CAGCTTAGCT AGCTTAGATG
68301 TTTGGCACCA GTTTGAATCA TGCTTTTAC AGCTGGCTCC ATGTAGTCTT
68351 TCCAAACATT TTGGCCTTTC CTGAGCAGCC CTTGTAGATA TTGTCTGTAT
68401 GATGCATTTT GACACAAGGT GATATTTTT GTGATATCAA AATTCCACAT
68451 TTACCCATTAA GAGTTACAGC CCTGGGGTTC ACAGTACCAA GGGGGACCCA
68501 GAGCCTCAGG ATTGGCCAGG CTCATTTGC CGTGGAGTAT CAGTTGTCT
68551 TGAAATTGTG GGGAAAAAATT CTAAGTTGAA TTCACTGGTA AGTAATTTT
68601 TAAAATTCA TAATGCAGAT TACATCCAAA ATTGATTAA AAAATTAAAA
68651 CATAAGACTG CAGAGAAATT CTGCATTCA ACTCCAATAC TATCCAGACT
68701 TCAGAAATAA CTTATCAGTT ATTTCTGTAA GCTTCTTGCT TACCTGGATA
68751 CCTGACAGGT GAGATGGCTG TAGCAGACAC TGGCAGTTCC CTGCCCACAC
68801 ACCTGTCCCT GTCCACAGCT GCACAAGCA GCTCTGTGTG CAATTGCCAG
68851 CATCTGCTCC TCTGTTCTA GGGATCTT GTAGAAAAA TGCTGCCATA
68901 TTGTTCTC ACCTATTAGT CTTGTCTCCC AGTCAAGAGA ATAAATTAT
68951 GCAAGCAGAG ATTGTACTTT ACAGTATTG GTCTTGAGC TTGGCATTAG
69001 GTTGCATTTG TAAAATGTG GCATGGCTTC CTCATCCCCC AATAGGAAC
69051 TTGCCAGCCCT TTGTTCTC ATGGAACCTC CTTTTTGAA AAGAGCACCA
69101 AAGGAGTAAA AATACTGTGG AGGGAGCAAC CCTCCTTGC CATATGCTCT
69151 CATTGGGAGA CATGTGGAGC AGTCTGAAGT CATTAGGCC ACTCTCTGGG
69201 AGAGCACATC CTATGATGTT CTCCAGCCT AGCCCCCTCC ACTGTGCTCA
69251 AGTCCAAGCT GACCAGCTT CTGACCACAG TGAAACAAA GATGATTGTC

69301 AGTGGGCCCC AGAATCCTAT ACCCAGA

FEATURES:

Start: 2132
Exon: 2132-2314
Intron: 2315-17055
Exon: 17056-17182
Intron: 17183-20983
Exon: 20984-21071
Intron: 21072-41719
Exon: 41720-41831
Intron: 41832-45391
Exon: 45392-45550
Intron: 45551-47878
Exon: 47879-48031
Intron: 48032-54612
Exon: 54613-54720
Intron: 54721-59290
Exon: 59291-59458
Intron: 59459-63791
Exon: 63792-63942
Intron: 63943-66164
Exon: 66165-66346
Stop: 66347

CHROMOSOME MAP POSITION:

Chromosome 1

ALLELIC VARIANTS (SNPs):

DNA

Position	Major	Minor	Domain
1722	G	C A	Beyond ORF(5')
1767	C	G A	Beyond ORF(5')
1840	C	G	Beyond ORF(5')
1857	T	G	Beyond ORF(5')
1945	G	T	Beyond ORF(5')
2007	A	C	Beyond ORF(5')
2769	C	G	Intron
3664	C	T	Intron
3827	G	A	Intron
4113	C	T	Intron
4337	A	G	Intron
4473	G	A	Intron
6455	T	G	Intron
6533	T	G A	Intron
6919	G	C	Intron
7305	G	A	Intron
7340	A	G	Intron
7466	A	G	Intron
7589	G	C	Intron
7810	A	C	Intron
9104	G	A	Intron
9503	A	T	Intron
9898	G	C	Intron
10196	T	C	Intron
12327	C	G A	Intron
13749	G	A	Intron
14150	T	C	Intron
14529	G	A	Intron
14653	G	A	Intron
15871	A	G	Intron
19244	G	A	Intron

19387	T	G	Intron
19447	C	G	Intron
20076	T	C	Intron
20492	T	-	Intron
20868	T	C	Intron
20941	T	C	Intron
21116	C	T	Intron
21701	G	A	Intron
21710	A	-	Intron
21826	C	T	Intron
21840	-	T	Intron
21841	-	C T	Intron
21843	-	C	Intron
22045	C	A T	Intron
22061	G	T	Intron
22348	-	A G	Intron
22682	A	G T	Intron
22783	-	T	Intron
23448	A	G	Intron
24960	G	A	Intron
24983	T	C	Intron
25390	T	C	Intron
26060	C	T	Intron
30245	C	G	Intron
33664	G	T	Intron
33883	C	A	Intron
34373	G	A	Intron
34558	G	T	Intron
43929	T	A	Intron
44309	T	- C	Intron
44997	T	G	Intron
46538	A	G	Intron
48153	T	C	Intron
48288	G	T	Intron
48412	G	A	Intron
48446	C	G	Intron
48456	G	C	Intron
48789	C	-	Intron
48859	G	C	Intron
49126	A	G	Intron
49378	T	G	Intron
49482	A	C	Intron
49741	G	A	Intron
49840	A	G	Intron
50102	G	A	Intron
50109	C	G T	Intron
50747	G	A	Intron
51272	G	A	Intron
52842	G	A	Intron
61837	A	G	Intron
62018	A	G	Intron
65562	A	G	Intron
65780	G	A	Intron
66092	G	A	Intron
66617	C	T	Beyond ORF(3')
66892	G	A	Beyond ORF(3')
67263	G	A	Beyond ORF(3')
67651	G	T	Beyond ORF(3')
67935	C	T	Beyond ORF(3')
69000	T	G	Beyond ORF(3')
69134	C	T	Beyond ORF(3')

FIGURE 3, page 24 of 42

Context:

DNA
Position
1722

TTGCCACGAGATGGCTGGATCTTCTGCAACAAATCCAGGAGTTCTCCTTTTG
TTTATAATTGCTCCAATAGATGTTAGGATTAACTCTGCTTTAAAGCAGAACATC
GCCATCCCAGGTGTGCAACCACGAAAAAATTAGACATCCGTGAGAGACAATGCCCTCCAT
GGCCAGTTCCAGGCAGAGAGAACAGCTCTGGCTGACGCCAAGGCTCCGGCCGAG
AGGTCTTAAGTGGAGTAACCAGTCTTCAGAACCCCCGCTCCAAGGCCACCGACGCGCTG
[G, C, A]
CGCTGCAGCCCTGGACCTGCTGGGGCCTTCTCGAACCGCATGCTGACAGCGGAG
TGGCAACTGGGAGGGTCAGAGGTGACCCCCGGTCCGACAGCACCTCCGAGACCCAGCTCCA
GCTCCCTACTCCGGCTCTGGAGGCGGGCCGGCAGTGCCGCGAGGCCAGCGCG
CGAGCTCTCCCCAGCAGCGGGACGCCAACCCCTGCGCGCGCGCGGGCTCGGTG
GGTCTCCGCTCTGCGCCCTGCGCGCCAGCCGACGGCCAAACGCT

1767

AGTTTCTCTTTTGTATAATTGCTCCAATAGATGTTAGGATTAACTCTGCT
TTTAAAGCAGAACCGCATCCCAGGTGTGCAACCACGAAAAAATTAGACATCCGTGAGA
GACAATGCCCTCATGGCCAGTTCCAGGCAGAGAGAACAGCTCTGGCTGACGCCA
AGGCTCGGGCCGAGAGGGTCTTAAGTGGAGTAACCAGTCTTCAGAACCCCCGCTCCAA
GCCACCGACGCGCTGACGCTGCAAGCCCTGGACCTGCTGGGGCCTTCTCGGACCCGC
[C, G, A]
TGCTGACAGGGACTGGCAACTGGCAGAGGTGACCCCCGGTCCGACAGCACCTCCC
GAGACCCAGTCCAGCTCCCTCACTTCCGGCTCTGGAGGCGGGCCGGCAGTGGCG
CCGAGGCCAGOGGGAGCTCTCCCGAGCAGCGGGACGCCAACCCCTGCGCGCC
GCGCGGCTGGTGGTCTCGCTCTGCGCCCTGCGCGCCAGCCGACCCCCGAC
GGCGCCCCAACGCTGTTGCGCGCCGCCCCAGCCGGCTCGCGCTGGTCCGG

1840

TCGCCATCCCAGGTGTGCAACCACGAAAAAATTAGACATCCGTGAGAGACAATGCCCTCC
ATGGCCCAGTTCCAGGCAGAGAGAACAGCTCTGGCTGACGCCAAGGCTCCGGCCCG
AGAGGGTCTTAAGTGGAGTAACCAGTCTCAAGACCCCGCTCCAAGGCCACCGACGCC
TGACGCTGCAGCCCTGGACCTGCTGGGGCCTCTCTCGGACCCGCTGACAGCGG
GACTGGCAACTGGCAGAGGTGACCCCCGGTCCGACAGCACCTCCGAGACCCAGCTC
[C, G]
CAGCTCCCTCACTTCCGGCTCTGGAGGCGGGCCGGCCAGTGGCCCGAGGCCAGCGC
GGCGAGCTCTCCCGAGCAGCGGGACGCCACCCCTGCGCGCCGCGGGCTCGGG
TGGGGTCTCGCTCTGCGCCCTGCGCGCCAGCCGACCCCGACGGCGCCCCAACG
CTGTTGCGCGCGCGCCGCCCCGCCCAGCCGGCTCGCGTGGCCGGTCTCGCCCGAG
CCCTCGATCTCCCGTACTTCTCGGCCAGGCCCTGCGCCCTCTGGGACCATGTTGCGC

1857

CAACCACGAAAAAATTAGACATCCGTGAGAGACAATGCCCTCCATGGCCAGTTCCAGG
CAGAGAGAACGAGCTCTGGCTGACGCCAAGGCTCCGGCCAGAGGGTCTTAAGTGG
AGTAACCAGTCTCAAGACCCCGCTCCAAGCCACCGACGCGCTGACGCTGAGCCCTGG
ACCTGCTGGGGCCTCTCTCGGACCCGCTGACAGCGGACTGCCAAGCCTCCAGCTCC
AGGTGACCCCCGGTCCGACAGCACCTCCGAGACCCAGCTCCAGCTCCCTCACTTCC
[T, G]
GCTCTCTGGAGGCGGGCCGGCCAGTGGCCGAGGCCAGCGCGAGCTCCCTCCCG
CAGCGGGGACGCCACCCCTGCGCGCCGGCTGGGTGGGCTCCGCTCTG
CGCCCTGCGCGCCGAGCCGACCCCGACGGGCCAACGCTGTTGCGCCGCGCC
CCGCCCCAGCCGGCCCTCGCGTGGTCCCGTCTGCCCCGAGCCCTCGATCTCCCGTGA
CTTCCTCGGCCAGGCCCTGCGCCCTCTGGGACCATGTTGCGCTGGCTGCCGGACTTCGT

1945

CAAGGCTCCGGCCCGAGAGGGTCTTAAGTGGAGTAACCAGTCTCAAGACCCCGCTCCC
AAGCCACCGACGCGCTGACGCTGCAAGCCCTGGACCTGCTGGGGCCTCTCCCTCGGACCC
GCATGCTGACAGCGGGACTGGCAACTGGCAGAGGTGACCCCCGGTCCGACAGCACCT
CCCGAGACCCAGCTCCAGCTCCCTCACTTCCGGCTCTGGAGGCGGGCCGGCAGTG
CCGCGAGGCCAGCGCGCGAGCTCCCTCCCCAGCAGCGGGAGGCCACACCCTGCGC
[G, T]
CCGCGCGGGCTGGGTGGGCTCCGCTCTGCGCCCTGCGCGCCAGCCGACCCCG
ACGGCGCCCAAACGCTGTTGCGCCGCGCCGGCCAGCCGGCTCGCGCTGGTCCC
GGTCTCGCCCCGAGCCCTCGATCTCCCGTGAATTCTCGTGGCCAGGCCGCTGCGCCTCT
GGGACCATGTTGCGCTGGCTGCCAGGACTTCGTGCTGCCACCGCGCCAGGACGCC
GAGCAGCCGACGCGCTACGAGACCCCTTCCAGGCAGTGGACCGCAATGGGACGGAGTG

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.10.485033; this version posted May 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

2007	GCCACCGACGCGCTGACGCTGCAGCCCTGGACCTGCTGGGGCCTTCTCGGACCCGC ATGCTGACAGCGGGACTGGCAACTGGCAGAGGTGACCCCCGGTCCGACAGCACCTCC CGAGACCCAGCTCCAGCTCCCTCACTTCCGGCTCTCTGGAGGCGGCCAGTGCC GCCGAGGCCAGCGCGAGCTCTCCCCAGCAGCGGGACGCCACACCCTGCGCG CGCGCGGGCTCGGGTGGGGCTCCGCTCGGCCCTGCGGCCAGCGCACCCCCGA [A, C] GGCGCCCCAACGCTGTCGCCGCGCCCCGCCAGCCCCGCCCTCGCGCTGGTCCCG TCTCGCCCCCGCAGCCCTCGATCTCCCGTACTTCTCGGCCAGGCCCTGCGCTCTGG GACCATGTTGCGCTGGCTGCGGACTTCGTGCTGCCACCGCGGCCCTGCAGGACGCGA GACGCCAGCGCTACGAGACCCCTTCAGGACTGGACCGCAATGGGACGGAGTGGT GGACATCGGCAGCTGCAGGAGGGCTCAGGAACCTGGCATCCCTCTGGCCAGGACGC
2769	TGGGGCCCGACCGGCACCCCGTAACAGAAGTGGGTATAATACGAAAGTCTACTGGT ATTGTCCAGATAAAATGAGTGTGGAACACTCTGGCCACGGGCACTGTTAAATT AAGACACTTTGTCTGAATCCATCCCAGGTTTTGTTCTGTTAAACCTTGCA ACATGTAATCCGTTTAGCTGTCAGACTTCAGTGGTCCCAAGTTGTATAAAGGGCGA CACATTGATCTCTTCGAAGCTGCTTGTACAGCAGCTATGTGTATTGTCTACTGTT [C, G] AAAAGTGGTAAACCAATCGCGTGTCCCCCACTCCCTGTTGAGAAGGAATGGCGGC ATTCCATTGTTAAAGACATTCTAGGTTAATGCCCTAGGTACATAAATTGATCTGAAGGG TTGACTTGACCTGCGACTGAGCAATTTCATTTCCTGAGTCATCTTAACGTGCCCCGT AACTCTGCCCTTTAGTAGGGTGGAGATATGTTGAACTCTCCAACCTGTTGAAGCGT TCCCTGACACTGGCATTCTTATCCAAAGAGGAAAGTGAATTAGTTACTATGAGGCC
3664	GCTGATTGTCACGAAATGGCCAGTTGGAGTTCCCCACCATGTCCAATCATTGGCTGGA AGCAGCCCAGGAAAGGGACGACCTTGCTGCAGTCAGCAGATGCCAGGGTTAGAGGC TAGAGAGTGAAGTCACAGTGTGTTCTCACAGTAGGTGCTTGAAGGGAGATCTCAGTG GTACAACCTCCATGGTCCCTACAATATACAAAAGCTTTGGAGTGCTCAATGTTAA GATTGTAAGGGATCCTGAGATCAAAAGCTTGTGAGATTGCTGCTGTATCACCATTAA [C, T] GTAACCTGCATCATATTCTGTTATATGTTGTCTAGTATATGTTACCAATTCTTTTA AATCACCTTTACTTTATTGATAGTTAAAAGATTGTAAGTGAATTGCAATGGATGT CCTTGATTCATTTCCTCATTCTGTCAGTTACTTCGTAGGATAAATTGAGGAGT GGACATTGCTGAGTCTGAAGGTAACACACATTAAACTGGGATACGTATTGCCCTTCGG AACACCTTAGACCCATTTCACTCTTGTACTGACAGTGCTTCTCCACATCCTCGCT
3827	GAAGGGAGATCTCAGGGTACAACCTCCATGGCCCTACAATATACAAAAGCTTTGGAG TGCTCAATGATTAAAGATTGTAAGGGATCTGAGATCAAAAGCTGAGAATTGCTG CTGTATCACCATTTCAGTAACGTCATCATATTCTGTTATATGTTGTCTAGTATA TGTTACCAATTCTTTAAATCACCTTTACTTTATTGATAGTTAAAACGATTGTAAG TGAAATTGCAATGGATGCTTGTATTCAATTCTGTCATTCTGGTCCAGTTACTTCGTA [G, A] GATAAAATTGAGGAGTGGACATTGCTGAGTCTGAAGGTAACACACATTAAACTGGGA TACGTATTGCCCTTCGGAAACCTTAGACCCATTTCAGTCTTGACTGACAGTGCTTGC TTCTCCACATCCTCGCTATTCAAGGGTATCAGTCTTGAAAGTCTCCATTCTGCAAGGT GAAATTCTTTCAATTCTGTCATTAGTCCATTAGTGTGCTATAGTGAATATCTGAG ACAGGGTAATTATAAAGAAAAGACATTATTAGCTCACAGTCCGCAGGCTGGGAAGT
4113	CAGTTACTTCGTTAGGATAAATTGAGGAGTGGACATTGCTGAGTCTGAAGGTAACACA CATTAAACTGGGATACGTATTGCCCTTCGGAAACCTTAGACCCATTTCAGTCTTTG ACTGACAGTGCTTCTCCACATCCTCGCTATTCAAGGGTATCAGTCTTGAAAGTC TCCTATTCTGCAAGGTAAATTCTTCTGTTAGTCCATTAGTGTGCTAT AGTGAATATCTGAGACAGGGTAATTATAAAGAAAAGACATTATTAGCTCACAGTTC [C, T] GCAGGGCTGGGAAGTTAAGAAGCGTGGTCTGGCATCTGCTGGACTCTGGGGAGGGCTT TCCTGCTGTGTCACAACATGGTGGAAAGTCAAAGTGGAAAGTGGACATGTGTGAAGAAGCA AAATCCGAGGGGTGCTTGGCTTATAGCAACCCAGCCTCGAGGGAACTGATCCATTACT GAGGGAACTAATTCACTGAGAGAGAGAACTCACTCACTACTGCAAGAATGACACC AAGCCATTGAGGGATGCTGCCCGTAACCTGACACCTCTGCTAGGTCCCTCC
4337	CATTTAGTGTGCTATAGTGGAAATATCTGAGACAGGGTAATTATAAAGAAAAGACATT ATTAGCTCACAGTCCGCAGGCTGGGAAGTTAAGAAGCGTGGTCTGGCATCTGCTGG

ACTCCTGGGAGGGCTTCCTGCTGTGTCACAAACATGGTGGAAAGTCAAAGTGGAAAGTGG
 ACATGTGTGAAGAAGCAAATCCGAGGGGTGTCCTGGCTTATAGCAACCCAGCCTCGAG
 GGAACGTACTGAGGGAACTAATTCACTGAGAGAGAGAGAACTCACTCACT
 [A, G]
 CTGCAAGAACATGACACCAAGCCATTCACTGAGGGATCTGCCTCCGTAACCCGACACCTCCT
 GCTAGGGTCCCTCCTCCAAACACGCCACATCAGGGATCAGACTCAACATGAGTTTGT
 GGGGACAAACAAAACGTAGCACTTGCTTTGCCTTGGTTCTATTCACATCCTCACAGG
 ATTGCATTATGCCTACCCATTGGTAGGGCAGTCTCTTTAATTGGTTACTGATTCAA
 ATGCTACCCCTCCAGAGACATCCTCACAGACACACCCAGAAATCATGTTTACAGTT

4473 TTCCTGCTGTGTCACAAACATGGTGGAAAGTCAAAGTGGACATGTGTGAAGAAC
 AAAATCCGAGGGGTGTCCTGGCTTATAGCAACCCAGCCTCGAGGGAACTGATCCATTAC
 TGAGGGAACTAATTCACTGAGAGAGAGAACTCACTCACTGCAAGAACATGACAC
 CAAGCCATTCACTGAGGGATCTGCCTCCGTAACCCGACACCTCCTGCTAGGTCCCTCCT
 CCAACACGCCACATCAGGGATCAGACTCAACATGAGTTTGTGGGACAAACAAA
 [G, A]
 TAGCACTTGCTTGCCTTGGTTCTATTACATCCTCACAGGATTGCAATTATGCCTAC
 CCATTGGTAGGGCAGTCTCTTTAATTGGTTACTGATTCAAATGCTACCCCTCCA
 GAGACATCCTCACAGACACACCCAGAAATCATGTTTACCACTGTTACAG
 GTCCAGACGAGTTGATAACATAAAATTACCATCACACATGGGATAGAATTAGGATTACAC
 AGTCAACCTTATGGGAGAAAATTTCAGAGGCATGTCAGGGTTATGTAATGTCAAGGA

6455 TGTGTTATTGCAATTGAGTGGAACTCAGGATTTCACTCCATTAGTAATTCTCTGTTAACAA
 AGAGGGTTCAATTTCATTTCATTTCATTAAATATTGCTTTTTTTCTGGAGAC
 AGAATCTTGCTCTATCACCAAGGCTGGAGTGCAGTGGTGCAGTCTCGGCTCACTGCAGCC
 TCTGCTTCCTGGATTCAAGCGATTCTGTGCCTCAGCCTCCAAGCAGCTGAGATTACAG
 GCACATGCCACCACACCTGGTTAACCTTGATTTCTAGTAGAGATGGGATTGCCAT
 [T, G]
 TTGGTCAGGCTGGTCTTGAATTCTGGCCTCTAGTGAATCTGCCTGCCTCTGCTCTGAAA
 GTGCTAAGATTACAGGCATGAGCTACCAGGCCATTCTTAATTTAATTGTT
 CAGACATGTTATGGTTCTGGCACAAATATTAAAGAAGACATGATATGAAATCACAGGGTGA
 ATTTTAGGGCATCACAACAGAAAGATTATGGTATAAGAAAACAATGGAATTCCAAC
 ATTTCTGTCAAATGTTCAAAATATATAAAATCTGTATCTTGTCTCTGATTT

6533 TTATTTCATTAATATTGCTTTTTTTCTGGAGACAGAACATCTGCTCTATCAC
 CAAGGCTGGAGTGCAGTGGTGCAGTCTCGGCTCACTGCAGCCTCTGCTCTGGATTCAA
 GCGATTCTTGTGCCTCAGCCTCCAAGCAGCTGAGATTACAGGCACATGCCACACACCT
 GTTTAACTTTGATTTCTAGTAGAGATGGGATTGCCATGTTGGTCAGGCTGGTCTT
 GAATTCTGGCCTCTAGTGAATCTGCCTGCCTCTGCTGAAAGTGTAAAGATTACAGGC
 [T, G, A]
 TGAGCTACCATGCCAGGCCATTCTCTTAATATTGTCAGACATGTTATGGTTTC
 TGGCACAAATATTAAAGAAGACATGATATGAAATCACAGGGTGAATTGGCATTACAAC
 AGAAAGATTATGGTATAAGAAAACAATGGAATTCCAACATCATTCTGTCAAATGTTCT
 AAAATATATAAAATCTGTATCTTGTCTCTGATTATATTCTAAATTGATGT
 TATCCTCTGCAAGAAATAAGTGTCTGAAAGAATGAAAAAAATGGAAGAACATTCTTAG

6919 ATGAAATCACAGGGTGAATTCTGGCATCACACAGAAAGATTATGGTATAAGAAA
 ATGGAATTCCAACTACATTCTGTCAAATGTTCTAAATATATAAAATCTGTATCTTT
 GTGTTCTCTCTGATTATATTCTAAATTGATGTTATCTCTGTCAGAAATAAGTG
 TCTGAAAGAATGAAAAAAATGGAAGAATTCTTAGTAAGGTATAAAATACCCTTCTATC
 TTTGTAGCATTCTAACGCCCTTGTCACTTCTGAAACTCCACATGCCATATTCCCTGA
 [G, C]
 TAGGCCACAGCCATGTACATTGATCCCTTATTTCTCTGCTGAGATTCTCTC
 ATTCCCCCTCTGCTGGTATATGATTGCCATTGTTAAGGCCAACTCACCTTCA
 TAATCTTCTAGCCCACCTCTTTATCGTATTCCAGAAAAACAAAAGCAGCTTCCACA
 AGACAAACATTCTGTAATACACTGCTTAACCTCTTTGACCTGCTGAGTTCAAAATCTT
 ATCTTTTAAGGATTGAATGGAGTCCACCAAGGTATCTATTGACAGGATTATGAAA

7305 GATTGCCATTGTTAAGGCCAACTCACCTTATAATCTTCTAGGCCACTTCTTAA
 TCGGTATTCCAGAAAAACAAAAGAAGCTTCCACAAGACAAACATTCTGTAATACACTGCT
 TAACTCTTTGACCTGCTGAGTTCAAAATCTTATCTTTAAGGATTGAATGGAGTC
 CACCAAGGTATCTATATTGACAGGATTATGAAAACAAAAGGATTGAGAAAGTT
 GAAGCCTAACTCTGAAACGTGGATCATAGTGTACTACACATTAACGTGTTAGTGGAT

[G, A]
TAATAGTTATTATAGGCTGTGGAAATCAGAACAGGGTTCAAATGTTTCACCGCTTGC
TAGACTGTGGCCTTGGCATGTTATTAAATGCCTGGAGGCCTCAAATGTTAAGTGGAAAT
GGTAAGACCTACCCAGTAACCTAGCATAAAATAGTAAATTCACTTAAATGTTTCAAA
CAGTGCCAGACATTGTTAATGAACCTGGGATATAGTGGTGAACAACACTGACAGCGTTC
TTCATTGTATTCTCAAAACCCCTCCCTATAGTAAGTAGGTCTGTGTGTAGGTGCA

7340 TAATCTTCCTAGCCCACTTCTTATCGGTATTCCAGAAAAAACAAAAGAAGCTTCCACA
AGACAAACATTCTGTAATACACTGCTAACCTCTTGCACCTGCTGAGTCAAAAATCTT
ATCTTTTAAGGATTGATGGAGTCACCAAGGTATCTATATTGACAGGATTATGAAA
ACAAAAGGATTGTTGAGAAAGTTGAAGCCTAACTCTGAAACGTGGCATAGTGTAA
CTACACATTAACGTGTTAGTGGATGTAATAGTATTATAGGCTGTGGAATCAGAAC
[A, G]
GGGTTCAAATGTTTCAACCGCTTGCAGACTGTGGCCTTGGCATGTTATTAAATGCCCTG
GAGGCCTCAAATGTTAACCTAGGAATGGTAAGACCTACCCAGTAACCTAGCATAAATAGTA
AATTCAATTAAATGTTTCAAACAGTGCCAGACATTGTTAATGAACCTGGGATATA
GTGGTGAACAAACACTGACAGCGTCTTCATTGTTACTCTCAAAACCCCTCCCTATAGTAAGT
AGGTCTGTGTGTGTAGGTGCATGGGAATAAAAATAAGCAAATAATGAACAG

7466 TTAAGGATTGAATGGAGTCCACCAAGGTATCTATATTGACAGGATTATGAAAACAAA
GGATTGTTGAGAAAGTTGAAGCCTAACTCTGAAACGTGGATCATAGTGTAACTACAC
ATTAACCTGTTAGTGGATGTAATAGTATTATAGGCTGTGGAATCAGAACAGGGTT
CAAATGTTTCAACCGCTTGCAGACTGTGGCCTTGGCATGTTATTAAATGCCCTGGAGGC
CTCAAATGTTAACCTAGGAATGGTAAGACCTACCCAGTAACCTAGCATAAATAGTAAATTC
[A, G]
TTCATTAAATGTTTCAAACAGTGCCAGACATTGTTAATGAACCTGGGATATAGTGGT
AACAAACACTGACAGCGTCTTCATTGTTACTCTCAAACCCCTCCCTATAGTAAGTAGGTCT
GTGTGTGTGTAGGTGCATGGGAATAAAAATAAGCAAATAATGAACAGGGTAAT
TTCAAAACAGCAGAAAGAGCTTCAACAAACTACCTGCCTTTATTAGATGAAACTCTC
AACTCTATGGTTGTTCTCCTGTCAATTCTGTTAAATGCTGTCAAGCCTGTTCCCTTA

7589 AACTGTTTAGTGGATGTAATAGTATTATAGGCTGTGGAATCAGAACAGGGTTCAA
ATGTTTCAACCGCTTGCAGACTGTGGCCTTGGCATGTTATTAAATGCTGGAGGCCTC
AAATGTTAACCTAGGAATGGTAAGACCTACCCAGTAACCTAGCATAAATAGTAAATTCATT
CATTAAATGTTTCAAACAGTGCCAGACATTGTTAATGAACCTGGGATATAGTGGTAA
CAACACTGACAGCGTCTTCATTGTTACTCTCAAACCCCTCCCTATAGTAAGTAGGTCTGT
[G, C]
TGTGTGTGTAGGTGCATGGGAATAAAAATAAGCAAATAATGAACAGGGTAATTTC
AAAAGCAGAAAGAGCTTCAACAAACTACCTGCCTTTATTAGATGAAACTCTCAAC
TCTATGGTTGTTCTCTGTCAATTCTGTTAAATGCTGTCAAGCCTGTTCCCTATCA
CCCTGGCCACGACTCTGTCTTCTGCTGGCTCTGTAGACTCTAACCCAAGGCTCATT
CTCTGCCTGGCTATCTGCCTCTGTGGCTCTTGCCACTACACTACATTCTGTGTTGCA

7810 CTGGGGATATAGTGGTGAACAAACACTGACAGCGTCTTCATTGTTCTCAAACCCCTCC
CTATAGTAAGTAGGTCTGTGTGTAGGTGCATGGGAATAAAAATAAGCAA
ATAATGAACAGGGTAATTCAAACAAAGCAGAAAGAGCTATTCAACAAACTACCTGCCTT
TATTAGATGAAACTCTCAACTCTATGGTTGTTCTCCTGTCAATTCTGTTAAATGCTG
TCAGCCTGTTCCATTACCCCTGGCACGACTCTGTCTTCTGCTGGCCTGTAG
[A, C]
CTCTAACCCAAGGCTCATTCTGCCTGGCTATCTGCCTCTGTGGCTCTTGCCACTAC
CTACATTTCTGTGTGACAGGGAGGACCATTCCCTGTGACCATAAAATTCTCTT
TGAAAAGAATCTCATCTGATTGGGACACAGCACATCTGTGAAACAGCATTAGACATTG
CCACTGCTCAGCAGCTGGGGAAATGTTACTGAGAAGCGTACAGTAGTTTTGAT
CTAACCATGGTGCACCCCTCCAGAGGGAAACCTATGAGTATTCAAGGACATGTGAT

9104 TTAAACGAATTATTGTTAGAACAGAAAAACAAATACTGTGTCTCATTTACAGGGGGAGC
TAAACCTGGTAAATGGGCATAAAAGATGGGAACAATAGACACTAGGGACTCCAAAAGG
GGGGAGGGAGGGAGGGCAAGGGCTGGAAAGCTCCACTGGGTACTTGTGTCACAC
CTGGGTGATGGCACGATTAGGAGCTCAAACCCAGTACACAGTACACCTGTAA
AGCTGATGGTGTAAACCCCTGAATCTCAAAATTATTAAATTTAAAAATCATTATA
[G, A]
GGATTGTTAAAAAGAAGGATTCTAGACAGGTGCAGCCAAACAAATTGTTTAAATGTTG
GCAGGCCGCCACCGCCAGTCACTTATGCTGCAATAGCCCAGTGTCCAAACATTCCAAACCT

	ACCTCTCTCCAAAAGAGAAGCTATACTTTCAGATGGCCCTGTGCTGGGTTCTCCCTGGAA GTTTCTGGGAAAGGGGCTTGAGTTGCCCGACTGGACTCTCCTGGAGTGGAGCCGGG GCTTCTGATCAGACGTGAGTGAGGCAGGAACCTCCCGGTCTCCCAGCGCAGCCCAGAGTG
9503	CATGTCCCAACATTCCCAACCTACTTCTCTCCAAAAGAGAAGCTATACTTTCAGATGGCC CTGTGCTGGGTTCTCCCTGGAAAGTTCTGGGAAAGGGGCTTGAGTTGCCCGACTGGAC TCTTCTGGAGTGGAGCCGGGCTCTGATCAGACGTGAGTGAGGCAGGAACCTCCCG TCTCCCAGCGCAGCCCAGGTGCGGCCCCACGCAAGTCCCGGGCTCGCGCTCGGCC TTTGCCTGAAGCCGTTAGGATGAGCCCTCTCCTCAGAGCTTAACCGATGAAGGTGC [A, T] TTGTGTTGGCCCTGAGGAGGATGCTGTCTAGGCCTCTTCCACTGGACGTGTG GTGGCAGAGATCCCGTCTCGTCGTCGACTTCACCCGCTGGGCTCACTCAGGCC GGAGCTGCAGGGAGACATCCTCGATGGACTCCCTCACGGAGATCTTTGGTACCTG GACTATAACAAGGATGGACCTTGGACATTGGTACAGGAGTGGAGGATGTA GGGCCATTCAATCTAGAGGAAGCGAAGGTGGTCACTGGGCTGTAATCAGAGAG
9898	ACCCCGCTGGGCTCACTCAGGCCGCGAGCTGCGAGGGAGACATCCTCGATGGACTCCC TCTACGGAGATCTTTGGTACCTGGACTATAACAAGGATGGACCTGGACATTTG AGCTTCAGGAAGGCTGGAGGATGAGGGCCATTCAATCTAGAGGAAGCGAAGGTGG GTCTCACTGGGCTGTAATCAGAGAGACGTTGGGCTGGAGCCCTGGAGAGGCATTGG CAGAGAGGGAAAATTACATGTTCAAGCTTGACCTGGGCCACTGCAGTGTTCAGGT [G, C] GTTGACCAGCGTTACCGTTATTAAGAATAACAACACAGCTAACACACATTCTCAAGTATT TTTCTCGTTTCTCCTGGCTGTAGTAAATCTCCAACCTCAGATTGCTCTCAAGATGT TGGCTACATAACGCCCTGTCTAGGAGTCACCTGTTCAATGTGCTCACCTGTCATTAGT CACCCAGAGGGCGTCTAGGCTAAAGATGCGCCCTCCCAGTCAGAGAACTGGAATAAT CACTCTACGTGTATTGGAGTGGGTGGTATTGGAAATTCTGATGTTATGTTGG
10196	GTGGTTGACCAGCGTTACCGTTATTAAGAATAACAACACAGCTAACACACATTCTCAAGT ATTTTCTCGTTTCTCCTGGCTGTAGTAAATCTCCAACCTCAGATTGCTCTCAAGA TGTTGGCTACATAACGCCCTGTCTAGGAGTCACCTGTTCAATGTGCTCACCTGTCATT AGTCACCCAGAGGGCGTCTAGGCTAAAGATGCGCCCTCCCAGTCAGAGAACTGGAAT AATCACTCTACGTGTATTGGAGTGGGTGGTATTGGAAATTCTGATGTTATGTT [T, C] GGTTCTGTTCTGGAGGGGCAGTGGAGTGGCTTTACTCTGGGTTCACTAGTGC TGAGGTTCTCTATAATATGCCCTAATTGATAGACCTAGTATCAGTACCGAGCTTAGG CTAACCCCTCTCTCCCAGAAGGCTAACCTACAGGCTCCTCTCAGCATGTTGTGCTTC GTACATACTCTATTGCACTATTCCAAGTCATTGGAATTATTATTGTATA TAATAATTACTTTATAAGTATTTGCTCTTGGATGTTGACCCGTAGACTGGAGAT
12327	GTCATGTTATTAATGCCCTGGAGGCCTCAAATGTTAACCTAGGTAATGGTAAGACCTACCC AGTAACCTAGCATAAAATAGTAAATTCAATTCAATGTTCAAACAGTGCCAGACATT GTTTAATGAACTGGGATATAGTGGTGAACAACACTGACAGGTTCTCATTGATTCTC AAAACCCCTCCCTATAGTAAAGTAGGCTGTGTGTAGGTGCATGGGAATAAAAAA TAATAAGCAATAATGAACAATAAAATTATTTATTAAAAAAAGAAATGATACTTAC [C, G, A] TTGTCGTGTTAAGATACAAAAGCAATAACTTTTATTGTGAAAATAGTGTGTTGGAC AATATATTGTTTGTGTTCTGTGAAAGTTGAGAAAACATAACAGAAGAGATAATG GTCAGACCTAAATAAAAGAACTTTGACTCAAATTACAGCAGTCTGCCAGAAAA CCAGCCCTTATCTAAAATAACAGACCAAGGAAACAGCCTGTTATGTCAGACTTATAGG AAGTCAGGTGCTATCTCTAGAGACAATACACAAAGCTATGCAATAACTGCTGTAACAGC
13749	TACAGGCCTGAGCCACCATGCCAGCCATAGACTATATATTTGATCTGATAACTGG TTCACTACTAAGTGAACACAGGAAGTAGCATCTATAGTGTGGATATGCTGGACAAAA GGACATTCCACTCTGGCAGGATGGCACAGAAATGTTGAGAGATTTATCATGCTACTCA GAATGGTGTCAATTAAAACCTATGAGTTGTTCTGGAGTTCCATTAAATAGT TCAGACCATGGATTGACCGCAGGTAACGAAACTGTGGAGAGTGAAACTGTGGATAAGGG [G, A] GGACTATTGTTAGTAAAGTCAGACTCATTAGGCAATCATAACTCTGATTGCCATCAG AAATGCTGAGAAATATGGTTAAAAAAACTGTTCAAAAATAGGGTCAAGGGATGTCCTT TAACCTGTTACTTCCAAATGTTAGTGAACAAACTGTGGCCCAAAGAGTGAAAGGAACAAA TGACTAAGAGAAAATCTGTTTCAAGGATGACAGATTAAAAAGAAGCAACTTGCTGAA CACTGAAAATCTCTCCACTGTAAGATAACACAAAATGGCTAAAATGGTTGGAATGAA

FIGURE 3, page 29 of 42

14150	ATAGGGTCAGGGATGTCTTTAACCTTGTACTTCAAAATGTTAGTGAAAACGTGCC CAAAGAGTGAAGGAACAAATGACTAAGAGAAAATCTGTTTCAGGATGACAGATTAA AAAGAAGCAACTGCTGAAACACTGAAAATCTCCACTTGTAAAGATAACACAAAACCTGG CTAAAACGTGGTGAATGAATATGGCCAACCTCAAGTCTGCACAGAACTAAGTGGTGTG TTACAGCCCCAATTTCACCACATATTTATACTAACTCCCCCGGATTTCACACATGA [T, C] CTGTGAGGTAGCATGAAGAGGTAACTATGCATGCCAAGGACTTGGGAGACCTCCCCATT TCCTCCACCAATCACCCACTAATCCCAGAATCCGCCCCAACCTTTCTAATAACTAC CTAAAGCCAGCATAGGGAGACAGATTGAGCTGGACTCCTGCTTGTGGTCACCT TGCATAAAAAGCTTCTTCAACACCTGGTATTAGTATTGACTTCTAGTC CGGGCAGCAAGCCCCTTTGGTCGGTACTATTCTGTCGTGATATTCCATTGGCA
14529	ACTAATCCCAGAATCCGCCCCAACCTTTCTAATAACTACCTTAAAGCCAGCATAGGG AGACAGATTGAGCTGGACTCCTGCTTGTGGTCACCTTGCATAAAAGCTTTTC TTTCTCAACACCTGGTATTAGTATTGACTTCTAGTCATCGGGCAGCAAGCCCC TGGTCGGTACTATTCTTGTGCTGATATTCCATTGGCCAAATATAAACCTCTTAGA TGAACATTCACTACGTAATGGGCCACAGAATGCTGTGACATTTCCTTGATTTA [G, A] CAGGTTACTTTACTGAATACCGTAGGCAGTTATAACACACTAAGTATTGTGATCTAA CATAGAAAAGATAACAGTAAAAATGGTAAATTTCACACTTGTGAGATTGGAG GGTATGTGCACATTGTTACAAGGGTATATTGCATGATGCTGAGGTTGGGTACAATTG AACCCGTGACCCAGGTAGTGAGCATAGTACCCATCGATAATTTCACCCCTGTCCA TTCCCTCCCCGTTCTGTAGTCCCAGTTCTGTTCTGTTCCATCTTATATCCGTGTGCA
14653	CTCAACACCTGGTATTAGTATTGACTTCTAGTCATGGCAGCAAGCCCC CGGTGACTATTCTTGTGCTGATATTCCATTGGCCAAATATAAACCTCTAGATGAA ACTTCAGTACGTAATGGGCCACAGAATGCTGTGACATTTCCTTGTGATTAGCAG GTTACTTTACTGAATACCGTAGGCAGTTATAACACACTAAGTATTGTGATCTAAACAT AGAAAAGATAACAGTAAAAATGGTAAATTTCACACTTGTGAGATTGGAGGGT [G, A] TGTGCACATTGTTACAAGGGTATATTGCATGATGCTGAGGTTGGGTACAATTGAA CTGTCACCCAGGTAGTGAGCATAGTACCCATCGATAATTTCACCCCTGTCCATT CTCCCCGTTCTGTAGTCCCAGTTCTGTTCTGTTCCATCTTATATCCGTGTGAC ATGTTTGCTCCATGTGTATGTGAGAACTTGTGGTGTGTTGTTCTATTCTGCGTT ATTGCTTAGGATAATGCCCTCAGCTGCATCCATGTTGCTGAGAGGACGTGATT TT
15871	AGGAGTTTATCAATTATTAGTCTTCAAAGAACCATCTTGGCTTGTAACTC CCAATGGTGTGTTTCTTCTATTACTTTGTCTTATTCTCAACTCTTT GCTTAATTAAAATAATTCTTGAGATTGAGATAAGCCTCAATGATGGTCACCGATT CCAGTCTTCTTCTTCTAATTATGCATTAAACCAGAAATCTTCTAAGTGTAGC TTAGTTGAGCTCACAGTTCAGATCTGTCTCAGTCTGGAGGTTGGAGATCTGACC [A, G] TGACCATGAAACCACATCCAGTCACAATGTGGCATTATTTTTAATT TGAGATAGAGTTTCACTTATTGCTTAGGCTGGTGTGCAATGGTGCATC AGCAACCTCACCTCCAGGTCAAGCATTCTTGCTCAGCCTCCAGTAGCTGG ATTACAGGATGCCACCATGCCAACATAATTGTATTAGTAGAGATGGGGT TCCATGTTGGTCAGGTTGGTCTTGAACCTCAGGTGATCCGCCACCTCAGCCT
19244	GTGGCATTATTGGTCATATTAGTCTTAACTC TGATCCTCATATTGGGATTCTGTATTGCAAATTGCTACTCAATAAAATT CCAAAGTAACCCAAAATATAACTCACAGTACTTCCAGGCATT GAGCAGTAAAAACTGAGTTGCTCAGCATGTACATTCTAGTAGAATAAGGCAAT ACTCTGCCTCTTGTTCAGCTCTACACTATTAACTAGCAAGTATCC [G, A] TTTGTCAGTTTGCATTGGTATTGCTGTTGGTAATT CCAAAGGTAGTGTGAAAGTGTCTAGTGTCTTAAGTGC TTATGGAGAAAATATGCGTGGATAAGCTTGC CAGCACACATTAAATGAGGTGC AATCAGTTGATGAAAGTGTGTAAT TCAGCATGTACATTCTAGCTAGTAGAATAAGGCA ACTCTGCCTCTTGTTCAGCTCAGGTGATCCGCCACCTCAGCCT
19387	CTCACAGTACTTCCCAGGCATT CATGGACATGCACAGAGCAGTGAAA TCAGCATGTACATTCTAGCTAGTAGAATAAGGCA ACTCTGCCTCTTGTTCAGCTCAGGTGATCCGCCACCTCAGCCT

	TCATACTATTAACAGCAAGTATCCCTTCAAGGTCTATTGTCGCCAGTTGCATT TTGTATTTGGTGGTAACTCCTTTAAAATGTCCTCAAAGGTAGTGCTGAAGTGCT GTCTAGTGTCCTAAGTGCAGAAAGCCATAGCATGCCTTATGGAGAAAATATGCGTT [T, G] GATAAGCTTGCCCCAAATTCAATGTTAGTGAATCACAGCACACATTAAAGGGTGC TTCAAACAGAACAGACATAAGACATGGTTATGTTAAATCAGTTGATGAAAGTGTGTA ATCAGAGGCTCACAGGAACCTAACCCCTGTTTCTGTAGGAACAATGGTTGGTATTG CTAATTCAAGTGTGCAATGAATATAGAACCTTATGGAAGATGATTGCTGTGAATAATGA GAATTAACCATACTCTTAAGAGTGCATTCTAAAGGAGAATATTCAAGAAGGGTATTG
19447	TCAGCATGTACATTCTAGTAGAATAAGGAAACTCTGCCTTCTGTTCAGCTC TCATACTATTAACAGCAAGTATCCCTTCAAGGTCTATTGTCGCCAGTTGCATT TTGTATTTGGTGGTAACTCCTTTAAAATGTCCTCAAAGGTAGTGCTGAAGTGCT GTCTAGTGTCCTAAGTGCAGAAAGCCATAGCATGCCTTATGGAGAAAATATGCGTT GGATAAGCTTGCCCCAAATTCAATGTTAGTGAATCACAGCACACATTAAAGGGTGC [C, G] TTCAAACAGAACAGACATAAGACATGGTTATGTTAAATCAGTTGATGAAAGTGTGTA ATCAGAGGCTCACAGGAACCTAACCCCTGTTTCTGTAGGAACAATGGTTGGTATTG CTAATTCAAGTGTGCAATGAATATAGAACCTTATGGAAGATGATTGCTGTGAATAATGA GAATTAACCATACTCTTAAGAGTGCATTCTAAAGGAGAATATTCAAGAAGGGTATTG CATATAATTCTTACTAACAGATGCTGCCTCACTGTCCTACATGGTCCAGATTCTCAT
20076	TCTCTCAGAACCTGTCATCTCCTCCAGGGTCCTTCTCCAAGAAAGTCTATCCTTCAC CACTAACAGTAATTGGCTTCCTCTTTCTGGAGAAGTCAGCTGTTATGCTGCTTC AGCACCAGACCCCTCTTACTTTGTTGTTCTTCTCATGTACAGTAGTCTAG GATTCTCATGAGCCTGTGAGCTGCTAGAAGGAATACAGCAGTGCCTACATTATTGCTT CTATTGTTATTCTATTCTCTTCGTCTGATTGTTCTCCTCTGTCCACAAACA [T, C] GCTCTAACTTCCCTAGTATTAAAAATTCTGTTTTGTTCTTTATCCTTGCTCC CTTATTGTTACTGCCAGATTGTTATTGTTATTGTTATTGTTGAGATGGAGTCTCACTC TGTCACCCAGGCTGGGTGCAGTGGCGCGATCTCAGCTCACTGCAACCTCCGCTCCAG CTTCAAGCAATTTCCTTTAGCTCCAAAGTAGCTGGGATTATGGCACCTGCCACC ATGCCTGGCTGATTGTTCTATTAGTAGAGACGGGGTTTACCATGTTGGCCACACTG
20492	CACTCTGTCAACCCAGGCTGGGTGCAGTGGCGCGATCTCAGCTCACTGCAACCTCCGCT CCCAGCTCAAGCAATTTCCTCTTTAGCCTCCAAGTAGCTGGGATTATGGCACCTG CCACCATGCTGGCTGATTGTTCTATTGTTAGTAGAGACGGGGTTTACCATGTTGGCCA CACTGCTCTAACTGCTGACCTCAGGTGAACCACCCGCTCAGCCTCCAAAGTGCTGG GATTGCAAGGTGTGAGTCACTGTCCTGGCTTTACTGCCAGATTAAAAGAATAGTC [T, -] GTGCTTAGCTCTATTCTCATTACTACTCTCTTTAACTCAGTCATATGATGTT TGCATAGTAATGCTAGTAATTAAAATGAGAAATAGTACTTTAAAATGAAT AGATCCTACTTTAATTGAATTATCTGGAGTTAGAATATCTTGATTGGATTAGT TGCTACTCTTAATTACATTACTGGTAAGGCCACTTGAGTCAGTCTTTGGAGGA ATATTATTATCTATAAGGCTGTTACAATTACTGAATTAAAATGTTATTGTT ATTTTATCTATAAGGCTGTTACAATTACTGAATTAAAATGTTATTGTT
20868	TAGTAATTTATTAAAAATGAGAAATAGGTACTTTAAAATGAATAGATCCTACTTTAAT TGAATTTATCTGGAGTTAGAATATCTGATTGGATTAGTTCTGCTACTCTTAATT ACATTACTGGTAAGGCCACTTGTGAAGTCAGTCTCTTGAGGAATATTATTATCTAT AAGGCTGTTACAATTACTGAATTAAAATGTTATTGTTATTGTTAATGTT CATTGTTAGTATTGATGTTGGGATAGGCATTTAAGCAAGTCTATAACTCACCTACATGCA [T, C] AATTGCTTAATCAGTTAAAGCTTCTCTAAATGAGAGATTGAAATTCTATAATT CTGTGGTTCTTATCAGTCTGAGTTTATTGCTTCTTATTGTTAAAGGAAAA ATTGAGGCTCAGAAATTGTCAGTCTCCAGACACTGGGTCTGACTATTCTGACCAA CAAGCAGAGTTGATTCTCAAAGGTAGCTCTCATGTTGGTCAACAATTGACTTTCACT TTAATATCCTGCACTAGAACACTCTGTTGTAAGTGTGGCTTAAAACACCTCCCTAGTC
20941	GAGTTAGAATATCTGATTGGATTAGTTCTGCTACTCTTAATTACATTACTGGTA AGGCCACTTGTGAAGTCAGTCTCTTGAGGAATATTATTCTATAAGGCTGTTACAA TTACTGAATTAAAATGTTATTGTTATTGTTAATGTTGTTACATTGTT GATGTTGGGATAGGCATTAGCAAGTCTATAACTCACCTACATGCAATTGCTT ATCAGTTAAAGCTTCTCTTAATGAGAGATTGAAATTCTATAATTCTGTTGCTT

FIGURE 3, page 31 of 42

	[T, C] CAGTTCTGAGTTTATTGGCCCTTTATTAAAGGAAAATTGAGGCTTCAG AAATTGTCAGTCTCTCAGACACTGGGTCTGACTATTCTGAACAACAAGCAGAGTGA TTCTCAAAGGTAAGCTTCATGTTGGTCAACAATTGACTTCACTTAATATCCTGCA TTAGAACTCTGTGTTGTAAGTGTGCTTAAACACCTCCAGTCATTGTCATTATGATA TCCAAGATCTTTGTCCTTCCATTCAATTGATGTGACATTATCTAAAG
21116	GTATTGATGTTGGGATAGGCATTAAAGCAAGTCATAACTCACCTACATGCATAATTG CCTTAATCAGTTAAAGCTTCTCTAAATGAGAGATTGAAATTCTAATTCTGTGGT TCTTATCAGTCTGAGTTATTGGCCCTTTATTAAAGGAAAATTGAGG CTTCAGAAATTGTCAGTCTCCAGACACTGGGTCTGACTATTCTGAACAACAAGCAG AGTTGATTCTCAAAGGTAAGCTTCATGTTGGTCAACAATTGACTTCACTTAATAT [C, T] CTGCATTAGAACACTGTGTTGTAAGTGTGGCTTAAACACCTCCCTAGTCATT GTATATCCAAGATCTTTGTCCTTCTCCATTCTTGTATGTGACATTATC TAAAGTGAAGATGGGAAGTGTAAAGCTCAGACTGGACTCTTCTTCAGGCCTCAAAG GATAGTGGAAATGGCAGGAAGTAAGGTTAACCTCATAGATGAGGAGCTGAAGAGTTG GTGTTGCTTTCTCCATTGATTCTAATGTGACAGTAAACTCATTGATTCAAACAA
21701	CATTGATTCAAACATAAGAAGACTAGCAGATTCATCACATTATTAAACCTAGATGTGACTG GAAAAAAAGGAAATTACTAAGCTCCAAGCTAACAAAGAAATACCTGTTAAACTTCA GAAAACAGAAATGCAAATTGAAACCTTATTGTCGGGCAATCAGTTGACTATTAAAGT CAGACTTTATACTCTTAAATGTTGTTCATGGGATAGAGCAGTAATCTGCAGGCCA GGTGCCTCAAATACTCTGTTGCTATAAACACAGGGCAGGAACGTGATTGATAAAC [G, A] TAAAACAGAAAAGGACAATTATATTGATTAATATTGTTGTAATATTTCAGTCCTCAC ATTGTCATAAAATCTTCTAAATGGCTTGTATTGAATTATCTCATTATATCTGTG CCAACAGCATTTCTCATCTTCTCATAATTCTTTACAAACAGCTGCTCAAAGAGGA AGGCTCAAAGTCTCAAGGCTGAGCACGTAATGACTTTGTTAGTACTAGATGAGAAGGGC TTTCCTGAGGAAATGAAAACATGAAAAGATAAACAGAATTGGACAGTGA
21710	AAACTAAGAAGACTAGCAGATTCATCACATTATTAAACCTAGATGTGACTGGAAAAAGG GAAATTACTAAGCTCCAAGCTAACAAAGAAATACCTGTTAAACTTCAGAAAACAGA AATGCAAATTGAAACCTTATTGTCGGGCAATCAGTTGACTATTAAAGTCAGACTTT ATACTCTTAAATGTTGTTCATGGGATAGAGCAGTAATCTGCAGCCCAGGTGCTCTC AAATACTCTGTTGCTATAAACACAGGGCAGGAACGTGATTGATAACGTAAAACAG [A, -] AAAGGACAATTATATTGATTAATATTGTTGTAATATTTCAGTCCTCACATTGCTAA AAATCTTCTAAATGGCTTGTATTGAATTATCTCATTATATCTGTGCAAACAGCA TTTCATCCTTCTCTCATAATTCTTTACAAACAGCTGCTCAAAGAGGAAGGCTCAA GTCTCAAGGCTGAGCACGTAATGACTTTGTTAGTACTAGATGAGAAGGGCTTCTGAG GAAATGAAAACCTAAAACATGAAAAGATAAACAGAATTGGACAGTGAGATATAAGAG
21826	CAGAAATGCAAATTGAAACCTTATTGTCGGGCAATCAGTTGACTATTAAAGTCAGAC TTTATACTCTTAAATGTTGTTCATGGGATAGAGCAGTAATCTCTGCAGCCCAGGTGC TCTCAAATACTCTGTTGCTATAAACACAGGGCAGGAACGTGATTGATAACGTAAA ACAGAAAAGGACAATTATATTGATTAATATTGTTGTAATATTTCAGTCCTCACATTG TCTAAAATCTTCTAAATGGCTTGTATTGAATTATCTCATTATATCTGTGCAA [C, T] AGCATTTCATCCTTCTCTCATAATTCTTTACAAACAGCTGCTCAAAGAGGAAGGCT CAAAGTCTCAAGGCTGAGCACGTAATGACTTTGTTAGTACTAGATGAGAAGGGCTTCC TGAGGAAATGAAAACATGAAAAGATAAACAGAATTGGACAGTGAGATAT AGAGCATATAATTCTGCTTCTAAAGTAATATTCTCTAGGAAAGTGAGGGCGTTCCC TGGCTGTTAGGCCAGAAATCATATTCTTATATTCTGTTAGCTTAGGAAATAATGCA
21840	TGAACCTTATTGTCGGGCAATCAGTTGACTATTAAAGTCAGACTTTATACTCTAA TGTGTTGTTCATGGGATAGAGCAGTAATCTCTGCAGCCCAGGTGCTCTAAATACTCTG TTGCTATAAACACAGGGCAGGAACGTGATTGTTATGATAACGTTAAACAGAAAAGGACAA TTATATTGATTAATATTGTTGTAATATTCTGCTCAAAGTCAAAGTCAGCATTGCTAAAAATCTTC TAAATGGCTTGTATTGAATTATCTCATTATATCTGTGCCAACAGCATTTCATCC [-, T] TTCTCTTCTAAATTCTTTACAAACAGCTGCTCAAAGAGGAAGGCTCAAAGTCAGGC TGAGCACGTAATGACTTTGTTAGTACTAGATGAGAAGGGCTTCTGAGGAAATGAAAA

CCTAAAACATGAAAAGAAGATAAACAGAATTGGACAGTGAGATATAGAGCATATAATAT
TCTGCTTCTAAAGTAATATTCTCTAGGAAAGTGAGGGCCTTCCTGGCTTAGGCCA
GAAATCATATTCTATATTTCTTGATAGCTTAGGAATAATGCAAATTCTAAGCCAA

21841 GAACCTTATTGTCTGGGCAATCAGTTGACTATTAAGTCAGACTTTACTCTTAAT
GTTTGTGTTICATGGGATAGAGCAGTAATCTCTGCAGCCCAGGTGCTCTCAAATACTCTGT
TGCTATAAACACAGGGCAGGAACGTGATTTTATGATAACGTAAAACAGAAAAGGACAAT
TATATTGTATTAATATTGTGTAATATTTCAGTCCTCACATTGCTAAAAATCTTCT
AAATGGCTTGTATTGAATTATCTCATTATCTGTGCCAACAGCATTCTCATCCT
[-, C, T]
TCTCTTCATAATTTCTTACAAACAGCTGCTCAAGAGGAAGGCTCAAAGTCTCAAGGCT
GAGCACGTAATGACTTTGTTAGTACTAGATGAGAAGGGCTTCCTGAGGAAATGAAAAC
CTAAAACATGAAAAGAAGATAAACAGAATTGGACAGTGAGATATAGAGCATATAATATT
CTGCTTCTAAAGTAATATTCTCTAGGAAAGTGAGGGCGTTCCCTGGCTGTTAGGCCAG
AAATCATATTCTATATTTCTTGATAGCTTAGGAATAATGCAAATTCTAAGCCAAAG

21843 ACCTTATTGTCTGGGCAATCAGTTGACTATTAAGTCAGACTTTACTCTTAATGT
TTTGTGTTCATGGGATAGAGCAGTAATCTCTGCAGCCCAGGTGCTCTCAAATACTCTGTTG
CTATAAACACAGGGCAGGAACGTGATTTTATGATAACGTAAAACAGAAAAGGACAATTA
TATTGTATTAATATTGTGTAATATTTCAGTCCTCACATTGCTAAAAATCTTCTAA
ATGGCTTGTATTGAATTATCTCATTATCTGTGCCAACAGCATTCTCATCCTT
[-, C]
TCTTCATAATTTCTTACAAACAGCTGCTCAAGAGGAAGGCTCAAAGTCTCAAGGCTGA
GCACGTAATGACTTTGTTAGTACTAGATGAGAAGGGCTTCCTGAGGAAATGAAAACCT
AAAACATGAAAAGAAGATAAACAGAATTGGACAGTGAGATATAGAGCATATAATATTCT
GCTTCTAAAGTAATATTCTCTAGGAAAGTGAGGGCGTTCCCTGGCTGTTAGGCCAGAA
ATCATATTCTATATTTCTTGATAGCTTAGGAATAATGCAAATTCTAAGCCAAAGCT

22045 ATATTTCAGTCCTCACATTGTCTAAAATCTTCTAAATGGCTTGTATTGAATTAT
CTCATTATCTGTGCCAACAGCATTCTCATCCTTCTCTCATAATTCTTTACAA
ACAGCTGCTCAAGAGGAAGGCTCAAAGTCTCAAGGCTGAGCACGTAATGACTTTGTTAG
TACTAGATGAGAAGGGCTTCCTGAGGAAATGAAAACCTAAAACATGAAAAGAAGATAAA
CAGAATTGGACAGTGAGATATAGAGCATATAATATTCTGCTCTAAAGTAATATTCTTC
[C, A, T]
AGGAAAAGTGAGGGCGTTCCCTGGCTGTTAGGCCAGAAATCATATTCTATATTTCTTT
GATAGCTTCTAGGAAATAATGCAAATTCTAAGGCCAAGCTTCAGAATAGACTAAGAAGTATT
AGCTTAGCTGCCATGACAAAATACCATAGGCTGGATGCATTAAACATGAAATTAGTT
TTTCACAGGCTGGGAGCTGGGAAGTTAAGATGAGAGTGCAGCATGGTTGGGTTGAG
TGAGGGCTCTTTCTGGCTTGCAGATAGACCCCTCTCACTGTATTGTATGGCAGA

22061 CATTGTCTAAAATCTTCTAAATGGCTTGTATTGAATTATCTCATTATATCTGT
GCCAACAGCATTCTCATCCTTCTCTCATAATTCTTACAAACAGCTGCTCAAGAGG
AAGGCTCAAAGTCTCAAGGCTGAGCACGTAATGACTTTGTTAGTACTAGATGAGAAGGG
CTTCTGAGGAAATGAAAACCTAAAACATGAAAAGAAGATAAACAGAATTGGACAGTG
AGATATAGAGCATATAATATTCTGCTCTAAAGTAATATTCTCTAGGAAAGTGAGGGCG
[G, T]
TTCCCTGGCTGTTAGGCCAGAAATCATATTCTATATTTCTTGATAGCTTAGGAATA
ATGCAAATTCTAAGGCCAAGCTTCAGAATAGACTAAGAACTATTAGCTAGCTGCCATGA
CAAATACCATAGGCTGGATGCATTAAACATGAAATTAGTTAGTTTACAGGTCTGGGA
GCTGGGAAGTTAAGATGAGAGTGCAGCATGGTGGGTTGAGTGAGGGCTCTTTCT
GGCTTGCAGATAGACCCCTCTCACTGTATTGTATGGCAGAGAGAGAGAGAGA

22348 GAAAAGTGAGGGCGTTCCCTGGCTGTTAGGCCAGAAATCATATTCTATATTTCTTG
TAGCTTCTAGGAAATAATGCAAATTCTAAGGCCAAGCTTCAGAATAGACTAAGAAGTATTAG
CTTAGCTGCCATGACAAAATACCATAGGCTGGATGCATTAAACATGAAATTAGTT
TCACAGGCTGGAGCTGGGAAGTTAAGATGAGAGTGCAGCATGGTGGGTTGAGTG
AGGGCTCTTTCTGGCTTGCAGATAGACCCCTCTCACTGTATTGTATGGCAGAGA
[-, A, G]
AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGGGATCTTCTCTTCTATTATAAGG
CCATAGTCTGTGGATCAGGGTTCCATTCTTATGACTTTATTGACTTACCCCCCTAA
GATGCTATCTCCAGATATAATCACACGGTGGGTAGGGCCTAACATTGGATTGGAG
GGACACAGCTCAGTCCATAGCAAAGGATAATGCAGAGGGTGGATATTAAAAGTAGCTA
CACAATTNTTAATATAATATTGAGATGGTAACCTTTTTTGAGATGGAGTCTAG

GGATATGCAATATAAAATATAAAATTGTGACACTGAAAATTAAAAATGGGAGGAGTGG
 GTAAAAGTACCTTCATATAACTTACTATTATATCCTCTTATTGAATTGACCCTTTATCA
 TTATATAGGAACCTTGTTCTCCTTACAACCTCTGACTTAAAGTTGTTTATATGATA
 [T, C]
 AAGTAAAGTACTCCTGCTCCTTGGTTCTGTTCCATGGAATATCTTTCCATT
 CTTCACCATCAGTCTGTTACAGATGAAATGAGTCTGTATGGGAGCATAT
 AGTTGGATCTAGTTTTAATCCACTCAGACACTGTGTTTTGATTGATAATTAAAT
 CCATTCACTGGTCAAGTAATTATTGATAAGTAAGGACTTTGACTACCATTTGCTTATT
 GTTTCATGGTCTTTATAGATCCTTATTCTTCTCCTCTTGCTGTCTTTTTT

 26060 GGTGTTGGTTGTGGTTACCAAGAGGTTACAAAAACATCTTAAGAGTTATAATAGTT
 ATTTTAACCTGATAACTTAATTTTATTGCAAAACCCCCAAAACAAAAAAATCTACAC
 TTTTACTTAATCCCCTGAAATTGGATTTGATGTCACAGTTACCTCTTCTATT
 GTGTATCCCTTAAATTATTGAGCTATTACTTTAAATAGTTCTCTTCTACTAC
 AGATGTAAGTGATTGCATACCATCATTACAGTATTATTGAAATTACCTGTGACTTT
 [C, T]
 TTTTATCAGGCCAGTTTATACTTCAGATGTTTGTGTTACTCATTAGCATCTTTCT
 TTCAGCTTGAGGAGCTCTTACGTTCTTAAATAGGTGCGGTATGATTATCTCC
 CTCAGCTATTGTTGTCTGGAAAGTATCTCTCTTCAATTCTGAAGGACACTTGCTGG
 GTACATTACCCCTGGTGGTATTTCCTCTGAACGTTAAATATCATCCCTTCT
 CTCCTGACCTGTTAGGCTCTGCTGACCAGTCTGTTCCAACCATAATTGGGACTGTCTTA

 30245 ATTTTAACCATCCATTGTTCTGCTCTAGATAACCCCTGACTAATATATAATTGGTAT
 GAAGTGATATCTCATGGCTTGATTATATTCTTCTCATGGTAGTGACTTTTTTGAC
 TTTGGGATATTGTTATTATTATTACTAGTGTGTTACTCTTCTCAGTAAAA
 GTGTTAGAACAAATTAAAGGCAGAATGTGACCAAGAGTTCCCTGAGTTATATAACCA
 TCATGGACCTCCCTCAAGTGTAAAGCATTAGTGTACTCATGTCACTCCAAATGTCA
 [C, G]
 TTGTTTCTCCATTCACTGTCCTTGTGTCCTAAACTGAAATTGATGGAAAAACAT
 CTGAATGGTCTTAATATGGTTGGATATTGTCCTCCAAATCTCATGTTGAAATATG
 ACCTCCAGTGTGGAAAGTAGGGACTACTGGGTACAGAGAGTGGATCCTCAATTAAATGGC
 TTGGAATAAGTGAACCTATTAGTCTCATGAAAGCTGGTTGTGATAAGAGCCTGGCATC
 TCATTCTCTGTCCTCTCACCATCTGACACACTTGCTCACCTTTTCTCAGCCA

 33664 TTCCAGAGTGTAGAAGTACACTGTCCTATCCTTCTAGGAGATCATTATAACACAAAAG
 CAGACAGTATATGAAACAGGGAAATTAGAGGCCAGATACCTATGACTTATATGAAAAAA
 TTTAAAGAAAATTAGCAAACCTGAATCAGCCATTAAATACACAAATCAATGC
 ATTCTATAAGAGCAGCTTAACAAAATTGTTAGAAGGCATTAAAGAAGACTCAGTATAGAA
 AAGATGTACTTCTCTCAAATTGGTGTAGAGATTCAATGCCATTAAAAACCCACCT
 [G, T]
 GTTTTTTGAGGAACTTGTCAGCTGAGTCTCAAATTATCAAAAGAGCAAAGGCCTAA
 GAATATCCAGGACATTCTGAAGAACTGTAAGGGAGCCAGGGCCTGCCCTATCAGATACC
 AAGGGTTGTATTAAGCCATAACCAAGTCAGTGTGTTCTACAGAAACAGACAAGTTAA
 CAAGTGAACATAATAGAGAGGCCAGAAACAGACCCATCCATTGGGATTGTCACGT
 GAAAGAAGTAGCTTGCAAAACTTGGAAAAGGAGAGTGTGTCATAAGATGATGCTCG

 33883 TAAAGAAGACTCAGTATAGAAAAGATGTACCTTCTCCTCAAATTGGTGTAGAGATTCAA
 TGCCATTAAAAACCCACCTGGTTTTGAGGAACCTGTCAGCTGAGTCTCAAATT
 ATATCAAAGAGCAAAGGCCCTAAGAATATCCAGGACATTCTGAAGAACTGTAAGGGAGCA
 GGGGCCTGCCCTATCAGATACCAAGGGTTGTTATTAAGCCATAACCAAGTCAGTGTGTT
 TCTACAGAAACAGACAAGTTAACAGTGAACACATAATAGAGAGGCCAGAAACAGACCCAT
 [C, A]
 CATATTGGATTGTCAGTGAAGAAGTAGCTTGCAGACAGACAAAAGGAAATTGGGATACCTGCC
 GTGTGCAATAGATGATGTCCTGCTCATGCAGACAAAAGGAAATTGGGATACCTGCC
 TTACCGTACACAAACACCAACCTAACAGTGAAGGTTAAACTATAACAGCTTGAGGTGGT
 GGGAGAAAATCTTATCTCAGTGTAGGGAAAGATTATTAAAAAGAAGACACAAA
 GCCCATACATAGGAATGAAAGATTGAATTGACGCTGCAATTAAAGATTAAATTGAC

 34373 TATCTTATCTCAGTGTAGGGAAAGAATTATTAAAAAGAAGACACAAAAGGCCATACA
 TAGGAATGAAAAGATTGAATTGACGCTGCAATTAAAGATTAAATTGACGCTGCGTAAAT
 CAAGAGCATTGACTTGGACAGCATAGAGTGGAAAGACAAGAGAAGGTATTGCCAGC
 TTATAACTGAAAGGATTAGAATGAATGATATAAGAAACTATGTAATAAGAAAAAGACAT
 ACAACCGGTAGAAAACGGGCAAAGACATGAACAGCATATTGACGTTGAAAGGAAACAGC

Sequence Data

	[G, A]
	GTTATACCCACAGCAAGACTATCTTATCTAGGAAGTTGTCAATACCCCTAAATGTTCTGT GGTTTAAGCTACAGAGTTGTAACTTCATTATTCAATAAAACTCAGTGGCAGGC ACTGTTTAGAACCTGGTTATAACTTGATGAAATTAAAAAAACTCCTGGCCTGTG GAGGATGCTATGTGTGGGAGTTGGTGGTGGGTCAAACAACAAATTACATTAAAATAG
34558	ACTTGAAGGATTAGAATGAATGATATAAGAACTATGTAATAAGAAAAAGACATACAAAC CGGTTAGAAAAACGGGCAAAGACATGAACAGCATATTTCACGTGAAGGAAACAGCGGTAG CAAATGAACATGGTAAGAGATGCTAACACAGTTAGTAATTGAAGGAAATGCAAGTTA TACCCACAGCAAGACTATCTTATCTAGGAAGTTGTCAATACCCCTAAATGTTCTGTGGTT TTAAGCTACAGAGTTGTAACTCATTATTCAATAAAACTCAGTGGCAGGCACTG [G, T] TTAGAAACCTGGTTATAACTTGATGAAATTAAAAAAACTCCTGGCCTGTGGAGGA TGCTTATGTTGGGAGTTGGTGGGTCAAACAACAAATTACATTAAAATAGAAAAT AGTACATAAAACCTATAAATATTGCAACCCAGAGTTATATTAAATGTAAGTAGT GACTAGGACTCTCATGAGATATACTCTGTGTGGACAAATGAAAGTTAAGTGTAA TCACATATGCAAGTCAAATAAAAGTGACACTAGAAAACACAATAATGAAATATCTGAA
43929	GGCATTAAAGTATTCTGCCATAGGGAAAGTGTAAAGTTGTAGGCTTTACTTTTATAGG TACTATATTGTCATAATAATCTCAGCACCTCATGGTGTCAAGGATCTGTGTCCCTGGTT GGTCAGATTATGTTATCTCTGGCATAAGGCACCTAACATATTCAATTAAAGGTTACAGA ATCTTTTGCTTCATGCTTAGCATTACACCAGTTGTTCCACCAAACCTTCAAA TTTGATTGTTCATTAATATTCTGCATACTGATGTAACCAAGTTCTATTATTGTGCAA [T, A] CTGCTCCTGAAACCCCTAGGAACCTCTGAAGGAGTTTATTATTGTTTTGTTTTGTT TTGTTTTGTTGTTGTTGAGACGGAGTCTGCTGTGCCCAGGCTAGAGTCAG TGGTGCAGATCTGGCTCTGCAAACCTGGCCTCCGGGTTACGCCATTCTCCTGCC AGCCACCGGAGTAGCTGGACTACAGGCACCCACCACTGCGCCTGGCTAATTGTTGTT ATTGTTAGTAGAGACGGGTTCACCGTGTAGCCAGGATGGCTCGATCTCCTGACCTT
44309	TTGAGACGGAGTCTGCTGTGCCCCAGGCTAGAGTCAGGGCGATCTCGGCTCTC TGCAAAACTCGGCCTCCGGGTTCACGCCATTCTCCTGCCTCAGCCACCGGAGTAGCTGG ACTACAGGCACCCACCACTGCGCCTGGCTAACATTGTTGTTAGTAGAGACGGGG TTTCACCGTGTAGCCAGGATGGTCTCGATCTCCTGACCTTGTAAATCCGCC TCCCAAAGTGTGGATTACAGGCCTGAGCCACTGTGCCGGCTTTTTTTTTTT [T, -, C] TTTATGGCTTGTCTTACACTTCAGATTGACTAAATTAAATATGCAATTAAATGAAGT CAGGAGTTACATTGCCACTAGTAACAAATGCCCTAACGTTACATAAGCATTATAAAATTG TTGGTATTAGTGCCTTCTCAGCTATGAGTATAAGATAATTACTAGTAGTTAGTT GCCTAGATAAAATTGTACACTATGTGAAGTTTATTACATAATTCTACGGTATTGTTA AGGTAGTTGATAACAGTTGAGACTACAATTGTATCTCATTGATAGTAAATGAA
44997	GAATTGTAAAAATATTATAGAATTGTTCTCAAACTATAGTAATGAGAATAGGT TGAAGGGGTGATGATTGAAACAATACCTCTCATTAGCTAAATTATAGAATCTAT TGCATGTTAAATGATAAGTCAGATTATAAAATATTGTTATAAACAGTAGGAAATGA GTTTAGGGTATTACACATACAGTTAATTGTTACATATTAAACATATCATGGT ATAAAATATGATGTGGATATAAAATTGAGATAAGGAAGTATTGTTAAGAATTGATGAAC [T, G] AATTCTTAAAGATGTCATCACCAAGTTGGTTCTAGCCTTATGAAAAATGGTGCAAT AAAAAAGATGACTATGATAAAATGTCGCCCTTCATTAAACCTAGACCAAGAGAAAAC ATACTGTGAATCTATGATGAATGAAAGAAAGTTGTAACTGTTGGTTGTATATTGTA TTACTGTTATTTCATTCTGTGAACGTACTGTACTTGTGTTGAGTAGACA ACTTATAATCTATGACTCAAATTGGTTAGTATAAAATTCTAGGAAATGAAAGTTCATATT
46538	TGTTATACTTATGGTCAACACTTTTATATTGTCGTAGATTCTGTACAAAAGATT TGACACTGTTAAGCCAGCATTCCTCAGAATGTACCCAAATCTCAAAATTATTAGG GGCAAAGCTAATGCTTAAAGAAAAAGGAGA [A, G] GGGATTGGGTGTGTTCTTAGGAACAGTAGTAACCTGACTTTAGAGAAACTGAAAT AAGCATTATTCTCTGGAAATTAGTTCTGCAAATTGAGGAGTTCCAAAGTCAACCTCAGGT TTGATACTCTAGAAAGACTCACATAACTCACTGAAAGCTTATTACCCCTGGTTATGG

	TTTATTACGGGAAAAGATCGGATGAAATCAGTCAGTAAAGAACATAGGCAGA
48153	TTATATCATCTGTTTATTTAGGTTACGGTTCAAAATCAGACAAAATGAACATAT TTGGTGGCTTCGACAGATGGAAAAGAAGGAGGTATCGCTCGCTTGAGGGAAATG GTACAAACGTCAAAATTGCTCTGAGACAGCTGTTAAATTCTGGCATATGAACAGG TAATTGTTATCACCCGTTGAATTATAACAAAGAGGAGTTAGTAAACGGATTCAATAAA TGTTAATGTATAATGTTGGATTCTGTTTAATACATGATAATCTTCACATATAC [T, C] CCATAAGGAGGATCACTTATAGGAGATTAGACTAAATAAAATCAGAGATTCTCATGACC AAGTTATGGGATTCTTAATTCTCATCATATTATTATAAAAGTTTTTTCTAAGTAGTT TTAAAGGAAGGGTAGAATTAGTTATTCTGAATCTGAGCAGAACGCACACT AACATAAGTTATGAAAGTGTACAATCTAACCTCTGAAAGGAAAATATAAGTTGAAG CCCTTGTGTAATTGACGTTGCTGAAATTGAGCTGAGTTGGAGTGACACCTCCATG
48288	AAATTGCTCTGAGACAGCTGTTAAATTCTGGGCATATGAACAGGTAAATTGTTATCACCC GTGGAATTTATAACAAAGAGGAGTTAGTAAACGGATTCAATAATGTTATGTATAATG CTTTGGGATTCTGTTTAATACATGATAATCTTCACATATACCCATAAGGAGGATC ACTTATAGGAGATTAGACTAAATAAAATCAGAGATTCTCATGACCAAGTTATGGGATT TTAATTCTCATCATATTATTATAAAAGTTTTTTCTAAGTAGTTCTAAAGGAAGGGTA [G, T] AATTTAGTTATTCTGAACTCTGAGCAGAACGCACACTAACATAAGTTTATG AAAGTGTACAATCTAACCTCTGAAAGGAAAATATAAGTTGAAGTCCTTGTGTAATT GACGTTGCTGAAATTGAGCTGAGTTGGAGTGACACCTCCATGAAGGCAGGGCGTGG CTTCTCCCCATGTAACCTCCAGCACCTAGACAGAGCTTGGCATGTGATAAGTTCAAGCGA GTGTTGAATGAGTCAATGAATGAACAAATGCATTACCTCTGAATCACTCTGTGCGC
48412	TGGGATTCTGTTTAATACATGATAATCTTCACATATACCCATAAGGAGGATCACTT ATAGGAGATTAGACTAAATAAAATCAGAGATTCTCATGACCAAGTTATGGGATTCTAA TTCATCATATTATTATAAAAGTTTTTTCTAAGTAGTTCTAAAGGAAGGGTAGAAT TTTAGTTATTCTGAAATCTGAGCAGAACGCACACTAACATAAGTTTATGAAA GTGTCACAATCTAACCTCTGAAAGGAAAATATAAGTTGAAGTCCTTGTGTAATTGAC [G, A] TTGCTGAAAAATTGAGCTGAGTTGGAGTGACACCTCCATGAAGGCAGGGCGTGGCTTC TTCCCCATGTAACCTCCAGCACCTAGACAGAGCTTGGCATGTGATAAGTTCAAGCGAGTGT TGAATGAGTCAATGAATGAACAAATGCATTACCTCTGAATCACTCTGTGCGCTTT GTTAACCTGGATTATTGAGCTATTGCTTCAGCTAACATGAAAGGGAAATACAG AGGTAAGTTAGAGTTGGGTCTTTATGGTCAATTAGCAGAACTGTCTAGTTGAGCA
48446	CATATACCCATAAGGAGGATCACTTATAGGAGATTAGACTAAATAAAATCAGAGATT TCATGACCAAGTTATGGGATTCTTAATTCTCATCATATTATTATAAAAGTTTTCTA AGTAGTTCTAAAGGAAGGGTAGAATTAGTTATTCTGAAATCTGAGCAGAACAG AGCACACTAACATAAGTTTATGAAAGTGTACAATCTAACCTCTGAAAGGAAAATATA AGTTGAAGTCCTTGTGTAATTGACGTTGCTGAAATTGAGCTGAGTTGGAGTGACA [C, G] CTCCATGAAGGCAGGGCGTGGCTTCTCCCCATGTAACCTCCAGCACCTAGACAGAGCTG GCATGTGATAAGTTCAAGCGAGTGTGAATGAGTCATGAATGAACAAATGCATTAC TCTGAATCACTCTGTGCGCTTGTAACTGGATTATTGAGCTATTGCTTCAGCC TAACTCAATGAAAGGGAAATACAGAGGTAAGTTAGAGTTGGGTCTTTATGGT CATTAGCAGAACTGTCTAGTTGAGCAGCCACAGATTATGTTCCATTATTATTCCATC
48456	ATAAGGAGGATCACTTATAGGAGATTAGACTAAATAAAATCAGAGATTCTCATGACCAA GTTATGGGATTCTTAATTCTCATCATATTATTATAAAAGTTTTCTAAGTAGTTCT AAAGGAAGGGTAGAATTAGTTATTCTGAAATCTGAGCAGAACAGCAGCACACTAA CATAAAGTTTATGAAAGTGTACAATCTAACCTCTGAAAGGAAAATATAAGTTGAAGTC CTTGTTGTAATTGACGTTGCTGAAATTGAGCTGAGTTGGAGTGACACCTCCATGAA [G, C] GCAGGGCGTGGCTTCTCCCCATGTAACCTCCAGCACCTAGACAGAGCTTGGCATGTGATA AGTTCAAGCGAGTGTGAATGAGTCATGAATGAACAAATGCATTACCTCTGAATCAC TCTCTGTGCGCTTGTAACTGGATTATTGAGCTATTGCTTCAGCTAACATCAATG TAAAGGGAAATACAGAGGTAAGTTAGAGTTGGGTCTTTATGGTCAATTAGCAGA ACTGTCTAGTTGAGCAGCCACAGATTATGTTCCATTATTATTCCATCATTGTTATC
48789	GCACCTAGACAGAGCTTGGCATGTGATAAGTTCAAGCGAGTGTGAATGAGTCATGAA

FIGURE 3, page 37 of 42

TGAACAAATGCATTACCTCTGAATCACTTCTGTCGGCTTTGTTAACGGATTATT
 TGAGCTATTGCTTCAGCTAACTCAATGTAAGGGAAATACAGAGGTAGTTAGAGT
 TTGGGTCTCTTATGGCATTAGCAGAACTGTCAGTTGAGCAGCCACAGATTATGTT
 TCCATTATTATTCCATATTGTTATCAAGGACTGTAAGGCCTTGAAATTCAACTCCC
 [C, -]
 CCCCCATAGTTTGTATTATTCCATGTAGATTAGATTATTCTGGAGAGTGTGTTGTT
 CTTGAGCAACAGAATACTCTTGAGAAGATTACGAAGTCCAGTGGTATCCTTCTTGCC
 TAGGAAATAGAGAAGAAAAAAAAAAAAAAATTAAAGAAAATCTAGTCTCCAGG
 ATTTAATTAGAACCTATCCTGGGAAGGCTATTTCCTTATATGAAGGTTGAAGATTC
 AAATCATGATTATAAGGGCTAATGTTGAGATAACCTTAGGTATTCTGACCACATACT

 48859 CATTACCTCTGAATCACTTCTGTCGGCTTTGTTAACGGATTATTGAGCTATTG
 CTCAGCCTAACTCAATGTAAGGGAAATACAGAGGTAGTTAGAGTTGGGTTCTC
 TTTATGGCATTAGCAGAACTGTCAGTTGAGCAGCCACAGATTATGTTTCCATTATT
 ATTCCATATTGTTATCAAGGACTGTAAGGGCTTGAAATTCAACTCCCCCCCCATAG
 TTTTGTTATTATTCCATGTAGATTAGATTATTCTGGAGAGTGTGTTGAGCAA
 [G, C]
 AGAATACTCTTGAGAAGATTACGAAGTCCAGTGGTATCCTTCTTGCTAGGAAATAG
 AGAACCAAAAAAAAAAAAAAAATTAAAGAAAATCTAGTCTCCAGGATTAAATTA
 GAACCTATCCTGGGAAGGCTATTTCCTTATATGAAGGTTGAAGATTCAAATCATGAT
 TATTAAGGGCTAATGTTGAGATAACCTTAGGTATTCTGACCACACACTGGATTATT
 GATAGGAAAGGCCACAGCTAAATAAAACTCAATGCAGTTATTCTAGTATGCAAG

 49126 GATTATTCTGGAGAGTGTGTTGAGCAACAGAATACTCTTGAGAAGATTACGAAG
 TCCAGTGGTATCCTTCTTGCTAGGAAATAGAGAAGCAAAAAAAAAAAAAAA
 ATTAAAGAAAATCTAGTCTCCAGGATTAAATTAGAACCTATCCTGGGAAGGCTATT
 CCTTATATGAAGGTTGAAGATTCAAATCATGATTATAAGGGCTAATGTTGAGATACC
 CTTAGGTTATTCTGACCACACACTGGATTATGATAGGAAAGCCACAGCCTAAATAA
 [A, G]
 TAAATACTCAATGCAGTTATTCACTGCAAGAAGTTGGTATTGAAAAAGTCAT
 GGGTATTGCAAGCAAATATGCACATTGCTTATGCCATTGTCAGATTCTACCTG
 ATACCACCAACAGGCATCCTCTGCTCTGTCACCCAGCTCCTGAGACCTCTTA
 TAGTATTGTAATTCTGCACACTAACCTCTAGACATGAAGAGAAAGCTGTACACAG
 TGTGGTGTAGTTCTATGGGCTCTGGACCTATGGTGTGTTCTCCTCTGCTGA

 49378 TGACCACATACTGGATTATGATAGGAAAGCCACAGCCTAAATAAAACTCAA
 TGCAGTTATTCACTGCAAGAAGTTGGTATTGAAAAAGTCATGGTATTGCAA
 GCAAAATATGCACATTGCTTATGCCATTGTCAGATTCTACCTGGATACCACCAAC
 AGGCATCCTCTGCTCTGTCACCCAGCTCCTGAGACCTCTTATAGTATTGTA
 TTCTGCACACTAACCTCTAGACATGAAGAGAAAGCTGTACACAGTGTGGTAGT
 [T, G]
 TTCTTATGGGCTCTGGACCTATGGTGTGTTCTCCTCTGCTGAAGGTCCATT
 CCCTCGGGGCTCTAAAGCCACCTTCTGTGACAAGCATATACTAACGATCTCAATCA
 AAGCCAGTCCCTCCCTGTCAGCCTCCCTCGAGTGTGCAATTGAGAACATAT
 TCATTGGATGATGGAAAACCCATTGTTTCCAGTGGATTGAAATTACTCGGGTAA
 TAGGCTGTATATATTCTCAAATTCCAGAGTATGTAACTAGGTCACTTTAGATT

 49482 TCCATGGTATTGCAAGCAAATATGCACATTGCTTATGCCATTGTCAGATTCTAC
 CTTGGATACCAACAGGCATCCTCTGCTCTGTCACCCAGCTCCTGAGACCT
 CTTTATAGTATTGATTCCTGCACACTAACCTCTAGACATGAAGAGAAAGCTGTCTA
 CACAGTGTGGTGTAGTTCTTATGGCTCTGGACCTATGGTGTGTTCTCCTCCCT
 GCTGAAGGTCCATTCACTCCCTGGGCTCTCTAAAGCCACCTCCTGACAAGCATAT
 [A, C]
 CTAAGCATCTCAATCAAAGCCAGTCTCCCTGTCAGCCTCCCTCGAGTGTGAATTG
 CAGAATATCCCATTGATGGATGATGGAAAACCCATTGTTTCCAGTGGATTGAA
 ATTACTCGGGTAAATAGGCTGTATATATTCTCAAATTCCAGAGTATGTAACTAGGT
 CACTTTAGATTCAAGATAGATTGTTCTGAAAGCTAGTACTTAGGAAACTAAGAA
 AAAGATCTTCAACCTGGTATGTAACCTGTCACATCAGTATGGGTAAACC

 49741 CTCGGGGCTCTCTAAAGCCACCTTCTGTGACAAGCATATACTAACGATCTCAATCAA
 GCCAGTTCTCCCTGTCAGCCTCCCTGAGTGTGCAATTGAGAACATATCCCATT
 ATTGGATGATGGAAAACCCATTGTTTCCAGTGGATTGAAATTACTCGGGTAAATA
 GGCTGTATATATTCTCAAATTCCAGAGTATGTAACTAGGTCACTTTAGATT

		GATTTTGTCTTGAATAGCTAGTACTTAGGAAACTAAGAAAAAGATCTTTCAACCTG [G, A] TATGTAGCTCTGTCAAACACATCATCAGTATGGGTAAACCTGTGTTCTGTGGTTGT CATTACCATAGTAGTGTATTGTATCATTGACAGTGTAAATAGTGTGGGTAGTGTCTTG TGGTTTCAGCTGCCACTCTGTACTGACTGCTTCCACTCCA
49840		ATCTTTCAACCTGGTATGTAGCTCTGTCAAACACATCATCAGTATGGGTAAACCTGTG TTCTCTGTGGGTGTCAATTACCATAGTAGTGTCAATTGTATCATTGACAGTGT [A, G] TAGTGTGGGTAGTGTCTTGTGGTTCAAGCTGCCACTCTGTACTGACTGCTTCCACTC CAACATCTTCTCTTATCTCAACACTGTAGGTCTACCTGTACTGTGTTCAAGCAT CTCTGCTTGCATGACCCAGGAGTGCCTCCACTCAATATGCCACCATGCATGGTCATCT TTCTGCTACTCCCTGTCTCCTGACCTGCTCCAGCAACACAGACAGACACCCTCCTCTT TCTATATGTATGTCATATGGGGAAATGCCCTTAGTACTACTCAGGAGTTAGTCCTCTGG
50102		CATTACCATAGTAGTGTATTGTATCATTGACAGTGTAAATAGTGTGGGTAGTGTCTTG TGGTTTCAGCTGCCACTCTGTACTGACTGCTTCCACTCCAACATCTTCTCTTATCTC AACACTGTAGGTCTACCTGTACTGTGTTCAAGCATCTGCTTGCATGACCCAGGA GTGCCTCCACTCAATATGCCACCATGCATGGTCATCTTCTGCTACTCCCTGTCTCCT GACCCTGCTCCAGCAACACAGACAGACACCCTCCTCTTCTATATGTATGGTGGGG [G, A] ATGCCCTTAGTACTACTCAGGAGTTAGTCCTCTGGGAAGCCTCTGTTCTAGTTCC TTTGTACAGCACTTCACATTGAATTCTGACGTTCTCTGTACTTATCTGTTGTGAG ACTGTGAGCTTCTTAGGCAGTAGCTACTGTATTCTAGCACCTGCCAGTGCAGGA AACCTTATAAGTAAATGAAAAGACAGAACTGACAGACTGGAATTAGAGCTCAAGCTG CCTCAATCTAAGCCATTAAGATGAAGGGAGCCGGCGTGGCTCACGCCTCTAAC
50109		ATAGTAGTGTATTGTATCATTGACAGTGTAAATAGTGTGGGTAGTGTCTTGTGGTTTC AGCTGCCACTCTGTACTGACTGCTTCCACTCCAACATCTTCTCTTATCTCAACACTG TAGGTCTACCTGTACTGTGTTCAAGCATCTGCTTGCATGACCCAGGAGTGCCTC CCACTCAATATGCCACCATGCATGGTCATCTTCTGCTACTCCCTGTCTGACCCCTG CTCCAGCAACACAGACAGACACCCTCCTCTTCTATATGTATGGTGGGGAAATGCC [C, G, T] TTAGTACTTACTCAGGAGTTAGTCCTCTGGGAAGCCTCTGTTCTAGTTCTTTGTT ACAGCACTTCACATTGAATTCTGACGTTCTCTGTACTTATCTGTTGTGAGACTGTGA GCTTCTTAGGCAGTAGCTACTGTATTCTAGCACCTGCCAGTGCAGGAAACCTT ATTAAGTAAATGAAAAGACAGAACTGACAGACTGGAATTAGAGCTCAAGCTGCCTCAAT CTCAAGCCATTAAGATGAAGGGAGCCGGCGTGGCTCACGCCTCTAAC
50747		CCAGCCTGGCAACGTGGCAAAACCCCATTCTACAAAAAATATAAAAATTAGTTGGACG TGGGGGTGTGCTGACTCAGGATGCTGAGGTGGAGGATCACTTGAGCTCGAGAGGC AGAGGTTGCACTGAGCTGGATCACACCATTGCAATCTAGCTGGTGTAGAATGAGAC CTTGTCTCAAAAAAAAAATAATAAAATAAAGGGAGATAAGGATTGGAAACAGAA GGAGCAGCATGTGGACAGAAATGTAGGCACAAGAAGGCATCACTCACTGAAGAGACTGAA [G, A] GTGGTTCACTGTGCCTCAAGACTGGTGGAGTGTGTTCCGAAAGATAATGATGAAAGAG CTGGACAGATAAACAGGGGCCAATGTAATAGGAGTCTGGATTTATTCTGAATATGGTA GGGGCTATTGTAGCATCTTATATAGGAAGTGAATGAGTACATTCACATTAAAGGAATA TCAACCTGAAAAAGAGTGGAGACATTGTTGGGGAGAGTGGAGTAGACTAGAGGCAGGG AGAATATTAAATAATTGAGGTAAGAAATGATGAACACCAGTATAAGGTGATGTCTTAA
51272		TAGACTAGAGGCAGGGAGAATATTAAATAATTGAGGTAAGAAATGATGAACACCAGTAT AAGGTGATGTTAAGGAATGGAGAAGGGAAATGAACTGAGAAATATTGGAAGTAGAA TCAACAGAAACTCACTGACTGGATATGGAGGTGAGAAAGAGAAGAGTCAGAATGAT ATTCTAATTCTAACATTGAGTGCATTCAAAGAGAAATACAATATCAGGTTCCATT GTGCATGCTGAGTTGAGATGTGAGGACATGTACAGGGAGCTGTCCAGTAAGCAATTGG [G, A] TATATCAGCTAGCCATTAAGAGAGAGATCTTGATAGAGAGGTTGTTGCTGAGTTGAGCC ATTGGAATGGCAGGATCACTCAAGAAGAGCTTATAATGAGAAGAAATTCTAGGAATAAG TCCAAAGGGAGAAGTAAAGAAGAAACTTGCAAGGACACTGAGAAGAAATAGCTCGAGG GATGGGAGAAAATCCAGAGAGAGGGATGGCATAGGAGTCAGTGGAGAAACGGTTTCAT GGGGGTCACTACTGGTAGTGAATATAAGAATATCTTGTAGGATTCTCAACCC

52842	TCAGGGTGGTTTGAGGGCTCAGTTAAGTCTCCTTAGGAAGGTCAGTCTGTAGCCTT GGCAAGTTACTAAAGCTCTGTGACTATTACCTCATCTAAGATGGGACTAAGCTTG GTGACATAGTTTACATACCAGGCACAGTGCCTGACTTTGGCTCTGCCTGAAGTCTT CCCTTGTATATGGTATGTTGGGAATAGGAGCCTCAAGCACTTATCCTTAAATATT TATCCTCCATCAGTCACTAACGTTACTCTGTACTIONTGTAGGTGCTGGGGTCCA [G, A] GGTATAAAAGGTACCTCAAAGTTACTGTTAAAGTGCAGGAAGGTTAAAGCAAATTAT GTTTAATGATTTGACAACTGACATGCAGGAAATTAAATAGGCCTATGCAGAAGAGGA GTTTATGTAACACTCTGTAGTTAGGAAACAGAGGCCCTGGAAGCAGTGAATCTCTGG GGAGGAATGCTGGTATTGGGAATCTCATGAAATGATAATATACTTAATTATCATG AGCAGCAAACACAGATTGCTAGGAGAAAGTCATCGTATGTTGCTGGCATTGGCACTTT
61837	GAGGAACCTCCATGTCATTTCATAGTAACTAGACCTTTGTTTTAACATTTCTAT CAATGTACACCAAGATTCAAATTCTCCATGTCCTCCCCAACACCATTAAGTGGGTGGT GGTCTACTACTATTGCTGTTGCTGTTATTCTCCCTCAGTCTGTAAAGTGTGTTGCT TCATATATTAGGAGCTTAATATTAGGTCCATATGAAGTTATAATTCTCTGGTAAAG TGACCCATTATCATTATGTAATGTCATTTGTCTCTGTGACAGTTGTGCTTAAA [A, G] TCTATTTGCTGATGTAATTATGCCACCCCTTTCTCTTGGGTTCCCGTTTTATGG AAATATCTTTCCATCCTTCACCTTCAGCTATGTTGCTGTTATTCACTCAGCAATTATATCTTTA TCATAGATAAGGTATAGTTGATTCTGTATGTTATTCACTCAGCAATTATATCTTTA GTTAGGGGATTTAATCCATTACATTAAAGCAGTTACTGATAGGAAGGACTTACTGTT GTCATTGGCTAGCTACCTTTATCTTGCTGTCCTGGCTTCTGTTCTGTTCCCTCCTC
62018	CATATATTTAGGAGCTTAATATTAGGTCCATATGAAGTTATAATTCTCTGGTAAAGT GACCCATTATCATTATGTAATGTCATCTTGCTCTTGACAGTTGTGCTTAAA TCTATTTGCTGATGTAATTATGCCACCCCTTTCTCTTGGGTTCCCGTTTTATGG AAATATCTTTCCATCCTTCACCTTCAGCTATGTTGCTGCTTAGATCTAAAGTGAGTC TCATAGATAAGGTATAGTTGATTCTGTATGTTATTCACTCAGCAATTATATCTTTA [A, G] TTAGGGGATTTAATCCATTACATTAAAGCAGTTACTGATAGGAAGGACTTACTGTT TCATTGGCTAGCTACCTTTTATCTTGCTGTCCTGGCTTCTGTTCTGTTCCCTCCT CTTCCTGGCTTCTCTGTGTTGATTTTTTTTGTTGAGTTGTTGAGTGTATGTTCTGAT TCCCTCTCATTTCCCTTGTGTCATTCTATAGATGCTATTGTTGTTTACATTGCA ACTACATAAAGCATACTAAAGTTATAGCAACTTATTTAAGCTGTTACAACCTAACCTC
65562	GAECTAAATTGAGACACATGCAGTCTGATTCTAACCCCTCTGTCTGCCAGCTCTGATCCA GAACTTTGCATGACTGATAGCCTGATAGATTGCTATGGCTGATAGACTGTCATTCTG ACCTAAAAGTCTGATCATTTACATCTGTCAGACATCTTGCAAGCCTTCGGTGTCACT TCCAAAGTTGTTAGTGGGAATTCAAGCCTTAAATAATCTAGCCCACCTTGTCACTC TCTGTGTAATAACCACATACAACAATTGGCTGCATCTCATAGCACATGGTACTCCTCCC [A, G] TTGTCTTGGTTGTGCCAGCAACACTGGTTTCGCTTCTCCCTGCTTGAGGTGAT TTCCAAGGCCAGGTCTTGTGCTTTTCCCAAGCTTCCCAGAGCTTCTCCATACTCCC CTTACTTCTGAGATTAACTGTTCTCTCTCAGCAGCTGTTGAGTTAGTAAGAAGGAGGCAGC AGCAGCACTGTTGGGGTGGAAAGTGTACCAAGCTTGGAGTCAGACCATTGGATCTCAG CCCTACCATTTCTACTTAGTATTAGGACAAATTCTCATTTCTAAGCCTCCA
65780	TCTAGCCCCACTTGTGCACTCTCTGTGTAATAACCACATACAACAATTGGCTGCATCTC CATAGCACATGGTACTCCTCCCGTTGCTTGCTGTCAGCAGCAACACTGTTTCTGCTTT CTCTTCCCTGCTTGTGAGGTCAATTCCAAGGCCAGGTCTTGTGCTTTTCCCAAGCTT CCCAGAGCTTCTCCATACTCCCCTACTTCTGAGATTAACTGTTCTCTCTCAGCGC TTGTCTAGTAAGAAGGAGGAGCAGCAGCAGCACTGTTGGGGTGGAAAGTGTACCAAGCTT [G, A] GAGTCAGACCATTGGATCTCAGCCCTACCAATTCTACTTAGATTAGGACAAATT TCTCCATCTTCTAAGCCTCAATTGCTCACTTACAAATTGATATAACATTACCTTGC AAAGATTGGTATGGAAGGTAATTAAACCCAGTATTAGAACATAGTAATTAAATAACTA TTATTACCATCATTACTATAGTTAGGACACTCACTGTTAGGTGCTATACAAAGAGGATCA TAAAAGGGATGTTGTCTGGCTCTGGAATAATGTTGCTTTACTGTATTAGA
66092	TTGGATCTCAGCCCTACCAATTCTACTTAGATTAGGACAAATTCTCCATCTT CTAAGCCTCAATTGCTCACTTACAAATTGATATAACATTACCTTGCAAGATTGGTAT GGAAGGTAATTAAACCCAGTATTAGAACATAGTAATTAAATAACTATTACCATC

bioRxiv preprint doi: <https://doi.org/10.1101/2023.09.11.553032>; this version posted September 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a [aCC-BY-ND 4.0 International license](https://creativecommons.org/licenses/by-nd/4.0/).

ATTACTATAGTTAGGACACTCACTGTTAGGTGCTACAAAGAGGATCATAAAAGGGATG
TTGTCTGGGCTTCTTGAATAAATGTTGCCTTACTGTATTAGAATATCATTCTG
[G, A]
GTCATAATTGTTGTTGTCATAATAATGAAACATACTGAATATTAAATTACCCCTTTT
TTTATTTTAGCCATGTTAGAAGGTCACAGCTGAATATGGTGGCCTCTTCGAC
GAATTATTCCAAAGAAGGAATACCAAGGACTTACAGAGGCATCACCCCAAACCTTCATGA
AGGTGCTCCCTGCTGTAGGCATCAGTTATGTGGTTATGAAATATGAAGCAAACCTTAG
GAGTAACCCAGAAATGATGTTGCATTTTGCTTACGCTGATAATTGAAACTTTCAACA

66617 ATGAAGCAAACTTGTAGGAGTAACCCAGAAATGATGTTGCATTTTGCTTACGCTGATA
ATTGAAACTTCACAACTCTGGAGTGACTTTTCTCCTCGAATTGAAACAAGTCTATG
GCAAAAGAAGCTGCATTTCACAAAAGGGAAAGATGGTACAATGGTCACTTCAAAC
TTTGGCCTAAATTATATGACACAGAAATGTTCAAATCATAGTTTAATGTGTTTGAA
AAGGCCACACAATTATACCTTATCTTCTTAAATCCTGAAATCTCTGCCCTGAATC
[C, T]
GAAATCTGAAATGTACTGGCTTGAACAAAATTGTTGTGTTAGAGTTATAAATCA
TTAACCTTATTCGGGGTTTACGTTATGCCAGTTCTTATATTAAATTCTTGT
TTTATATATTTGAATGTCTTATAGATTCTTAAATTCTTATAGAACCTTAATAG
AAAATCATTACATTAAATATACCTTACAGCAAAGCATCCAATAAGTATAGGGTTA
TGTCTTATTTCTTCAGCTGAATACGAATGAGCACAGTGGTGAATTCTGAAGGGA

66892 ATCCTGCAAATCTCTGCCCTGAATCCGAAATCTGAAAATGACTGGCTTGAACAAAATT
GTTTGTGTGTTAGAGTTATAAATCATTAATCTTATTCGGGGTTTACGTTTATGCC
AGTTCTTATATTAAATTCTTGTGTTATATTTGAATGTCTTATAGATTCTT
AAATTCTTATAGAACCTTAATAGAAAATCATTACATTAAATACCTTACAGCAA
AAGCATCCAATAAGTATAGGGTTATGTCCTTATTTCTTCAGCTGAATACGAATGA
[G, A]
CACAGTGGTGAATTCTGAAGGGAAAGTGATGAAATTATTTATTCAGTGGCAGCTT
TCCATTCTTACCACTGTACCATTATTGGTCTGGAGTTATACACTAATTTCAGTATAT
TACTGTTAAATTACCAACACAAGGCAATTATTGAAAGATCCGTTATCCTGCCATTG
CTTGAAAACCAGCAGGAAACGAAATCCTTGACTGTATCAGCTCTGCAGAGCATT
TGTCTTCTTGTCTTGTGTTCTACCTTGAATCAGATTCCGTTAGTCAGGAAGA

67263 CACTGTACCAATTGGTCTGGAGTTATACACTAATTTCAGTATATTACTGTTAAA
TTACCAACAAAGCAATTATTGAAAGATTCCGTTATCTGCCATTGCTTGAAG
CAGCAGGAAACGAAATCCTTGACTGTATCAGCTCTGCAGAGCATTGTTTCTT
TGCTTCTTGTCTTGTGTTCTACCTTGAATCAGATTCCGTTAGTCAGGAAGACTTCTGG
CCATTCTTAGTAACCTGAAATTCTTAAATTGATGATGAGTGGATTGATCATGAGCAA
[G, A]
TGATGTGCTTATTCCTCCCTCACTGTTGAATATCTTGAACCTGCTGTTCAATATGG
CAGCACAAGGTGAGAGATAACATATTAAATAGTAGTATGTTACTCTTACATTAGATA
CCTATATTAAATGAAAGGCCAATTGTAACATATAACATCATATTCTCTTGC
AAGTTTTAGGAACATGTTAGGATATAGGAGACTTAATTATAATGAGAGCATT
TATTGTTACTAAAGCCATTGTTATAGTCACATCTTCTTATTGTTGATTAGAACCT

67651 ATAGTAGTATGTTACTCTTACATTAGATACCTATTTAAATGAAAGGCCAATT
GTAAACATATACATTCAATTCTCTTGTGCCCAAGTTTAGGAACATGTTAGGATATAG
GAGACTTAATTATAATGAGAGCATTTTTATTTACTAAAGCCATTTTATAGTC
AACTATCTTCTTGTGATTAGAACTTAGAAAATTTACTAGTTGAAGTTAT
TATCAGTTTAATTAGTCTTAAACTCATTCACTTAATAATTCTGTTATAAATT
[G, T]
CCAGCATTTAAATGAAATCTAATGATGTAATAGGCATTCTTATTTGAACCTACCTC
TTTATTTCTGAACCAAAGAGAAAGATGGACTGGTGTGAAACATTAAATG
TAGTTTCTTAAATTAGTGTGTTGATAATGTCAGTATTGTTATAATGATAAG
CCTGGGATTCTACTTTAGGGTTATTGACTTTGAGTAATATAAAGTGACAATT
AAGGTACATGATCAGCTTTCTATTTACTCGTAAAATTATGAAATGAATAATT

67935 ATTCTGTTATAAATTGCCAGCATTTAAATGAAATCTAATGATGTAATAGGCATTCT
TTATTTGAACCTACCTCTTATTTCTGAACCAAAGAGAAAGATGGACTGGTGTG
AAACATTTTAAATGATGAGCTGGATTCTACTTTAGGGTTATTGACTTTGAGTAATA
TATAAAGTGACAATTAGGTACATGATCAGCTTTCTATTTACTCGTAAAATT
[C, T]

FIGURE 3, page 41 of 42

GGAAATGAATAATTTGCTAACAACTTGAAATTCAAACCTCTGGAAAATATGAAAATA
TTCATTGTTCAATTGAAATTAAATTGTAAGGTATGAATGTGATTGTCGTACATCTG
TATCTTTCCAAAAAATGATTCTGTATCTTGGAAAAAGCCGAGAGTTGAAGATAGTA
TATTCTGGTAGTACTGAATATTACAGTTCTATCAAAAATATATTGTTCT
AAAATTACTGTTTCCAGTTTATTTAGAGAAAATTCTTAAGTCTCAGTTCC

69000 TTCAGAAAATAACTTATCAGTTATTCTGTAAGCTCTGTTACCTGGATACCTGACAGG
TGAGATGGCTGTAGCAGACACTGGCAGTCCCTGCCACACACCTGCCCCACAGC
TGCACAAGGCAGCTCTGTGTGCAATTGCCAGCATCTGCTCCTGTCTCAGGGAACTT
TGGTAGAAAATGCTGCCATATTGTTCTCACCTATTAGCTTGTCTCCCAGTCAGAG
AATAAATTATGCAAGCAGAGATTGTACTTACAGTATTGTCTTGTGAGCTGGCATT
[T, G]

GTTGCATTTGTAAAAATGTGGCATGGCTCCTCATCCCCAATAGGAACCTTGCCAGCCC
TTTGTCTCATGAACTCCTTTTGAAAAGAGCACCAAGGAGTAAAATACTGTGG
AGGGAGCAACCCCTCCTTGCCATATGCTCTCATGGGAGACATGTGGAGCAGTCTGAAGT
CATTTAGGCCACTCTCTGGGAGAGCACATCCTATGATGTTCTCCAGCCTAGCCCCTTCC
ACTGTGCTCAAGTCCAAGCTGACCAGCTTCTGACCACAGTGTAAACAAAGATGATTGTC

69134 CTGTGTGCAATTGCCAGCATCTGCTCCTCTGTTCTCAGGGAACTTTGTTAGAAAATGC
TGCCATATTGTTCTCACCTATTAGCTTGTCTCCAGTCAGAGAATAAATTATGCA
AGCAGAGATTGTACTTACAGTATTGTCTTGAGCTTGCATTAGGTTGCATTGTAA
AAATGTGGCATGGCTCCTCATCCCCAATAGGAACCTTGCCAGCCCTTTGTTCTCATG
GAACCTCTTTTGAAAAGAGCACCAAGGAGTAAAATACTGTGGAGGGAGCAACCC
[C, T]
CTTGCCATATGCTCTCATGGGAGACATGTGGAGCAGTCTGAAGTCATTAGGCCACTC
TCTGGGAGAGCACATCCTATGATGTTCTCCAGCCTAGCCCCTTCCACTGTGCTCAAGTC
CAAGCTGACCAGCTTCTGACCACAGTGTAAACAAAGATGATTGTCAGTGGCCCCAGAA
TCCTATACCCAGA